AD-A087 829 TENNESSEE UNIV KNOXVILLE DEPT OF PSYCHOLOGY F/G 12/1
IS BAYESIAN ESTIMATION PROPER FOR ESTIMATING THE INDIVIDUAL'S A--ETC(U)
JUL 80 F SAMEJIMA N00014-77-C-0360

UNCLASSIFIED NL

1 OF 2
AD A
087829

LEVEL II

RESEARCH REPORT 80-3

12

IS BAYESIAN ESTIMATION PROPER FOR ESTIMATING THE INDIVIDUAL'S ABILITY?

ADA087829

FUMIKO SAMEJIMA

DEPARTMENT OF PSYCHOLOGY
UNIVERSITY OF TENNESSEE
KNOXVILLE, TENN. 37916

JULY, 1980

Prepared under the contract number NOOO14-77-C-360,
NR 150-402 with the
Personnel and Training Research Programs
Psychological Sciences Division
Office of Naval Research

80 8 11 005

(12) 152

(16) RR04204

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER: (9) Research Report 80-3	2. GOVT ACCESSION NO.: AD-A087829	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle): Is Bayesian Estimation Proper for Estimating the Individual's Ability?		5. TYPE OF REPORT & PERIOD COVERED: Technical Report
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s): (10) Dr. Fumiko Samejima		8. CONTRACT OR GRANT NUMBER(s): (15) N00014-77-C-0360
9. PERFORMING ORGANIZATION NAME AND ADDRESS: Department of Psychology, University of Tennessee, Knoxville, Tennessee 37916		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS: PE: 61153N; PROJ: RR 042-04; (17) TA: RR 042 04 01; WU: NR 150-402
11. CONTROLLING OFFICE NAME AND ADDRESS: Personnel and Training Research Programs, Office of Naval Research (Code 458), Arlington, VA 22217		12. REPORT DATE: 15 July 80
		13. NUMBER OF PAGES: 148
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office): (14) RR-1-3		15. SECURITY CLASS. (of this report)
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited. Reproduction in whole or in part is permitted for any purpose of the United States government.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Operating Characteristic Estimation
Tailored Testing
Latent Trait Theory

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

(Please see reverse side)

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

S/N 0102-LF-014-6601

Unclassified 404 32

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

There is a widespread belief among psychologists in the area of applied measurement that Bayesian estimation is better than the maximum likelihood estimation because of the additional information, i.e., prior. For example, many researchers in the computerized adaptive testing use Bayesian methods in the estimation of the examinee's ability. In this paper, this myth is debated theoretically, and in relation with the behavioral reality. Simulation studies are also used to show how biases caused by priors will affect the resultant estimation of the examinee's ability.

IS BAYESIAN ESTIMATION PROPER FOR ESTIMATING THE INDIVIDUAL'S ABILITY?

ABSTRACT

There is a widespread belief among psychologists in the area of applied measurement that Bayesian estimation is better than the maximum likelihood estimation because of the additional information, i.e., prior. For example, many researchers in the computerized adaptive testing use Bayesian methods in the estimation of the examinee's ability. In this paper, this myth is debated theoretically, and in relation with the behavioral reality. Simulation studies are also used to show how biases caused by priors will affect the resultant estimation of the examinee's ability.

The research was conducted at the principal investigator's laboratory, 409 Austin Peay Hall, Department of Psychology, University of Tennessee, Knoxville, Tennessee. Those who worked in the laboratory and helped the author in various ways for this research include Paul S. Changas, Dete Furlan, C. I. Bonnie Chen, Pamela Welch, Chi-Lin Tom and Robert L. Trestman.

TABLE OF CONTENTS

I Introduction

The main characteristic which distinguishes Bayesian estimation from the maximum likelihood estimation is that in Bayesian estimation we use the information given by the prior in one way or another, in addition to the information which is obtained directly from the set of observations. It appears to be a common belief among many researchers who are engaged in ability measurement that Bayesian estimation is superior to the maximum likelihood estimation, by virtue of this additional resource of information, the prior. In the area of computerized adaptive testing, for example, many researchers have used, and are using, Owen's method of Bayesian estimation (Owen, 1975), in order to accurately estimate the ability level of an individual.

It appears only logical to correct this common belief, however, and to say that the additional resource of information is valuable and desirable, <u>only if it provides us with a right kind of information</u>. If this is not the case, the additional resource of information is nothing but an obstacle, which may contaminate the estimation and lead us to biases, inefficiencies, and many other undesirable characteristics. We must pay our attention to this possibility, since not only such a resource of information will create contradictions in theory, but it may lead to serious social issues, such as the unfairness in personnel selection, etc. The objectivity of testing can be phrased in the principle of treating all the individuals of the same level of ability fairly and equally. If some statistical theory fails in fulfilling the requirement of this principle, then the resultant social issues are originated in the theory itself.

In the present paper, the effect of priors in Bayesian estimation will be considered, mainly from the standpoint of objective testing, which is closely related with the unbiasedness of estimation.

II Bayesian Estimators Vs. Maximum Likelihood Estimator

The main characteristic which distinguishes Bayesian estimation from the maximum likelihood estimation is that the former uses the prior as a part of the observation upon which the estimation is made, whereas the latter does not. In the estimation of a parameter which belongs to an individual, this prior is, in most cases, the density function of the population to which the individual belongs. We can say, therefore, that the maximum likelihood estimation is a population-free estimation, while Bayesian estimation is not.

In estimating the examinee's ability θ , the maximum likelihood estimator is the point of θ which maximizes the likelihood function, $L(\theta)$. When the estimation is based upon the response pattern, V , by virtue of the local independence (Lord and Novick, 1968), we can write for the likelihood function $L(\theta)$

$$(2.1) \qquad L(\theta) = P_V(\theta) = \prod_{x_g \varepsilon V} P_{x_g}(\theta) \quad ,$$

where $P_{x_g}(\theta)$ is the operating characteristic of the item response x_g to item g , or the conditional probability with which the examinee obtains the item score x_g , given ability θ , and $P_V(\theta)$ is the operating characteristic of the response pattern V .

It is well-known that the maximum likelihood estimator is asymptotically unbiased and normally distributed, when observations are taken from identical distributions (e.g., Kendall & Stuart, 1961). This implies that, when the test consists of n equivalent terms, or test items each of which has an identical set of operating characteristics, $P_{x_g}(\theta)$, of the items scores, x_g $(=0,1,\ldots,m_g)$, the maximum likelihood estimate is expected to be, approximately, equal to ability θ itself, if the number of items, n, is large enough and the amount of test information, $I(\theta)$, is substantially large. This characteristic of the maximum likelihood estimate also exists in a more general situation where the test items are not equivalent (cf. Samejima, 1975). It has been shown (Samejima, 1975, 1977a, 1977b) that this property of asymptotic unbiasedness and normality of the maximum likelihood estimate provides us with a good approximation even when the number of test items is relatively small, and is a useful characteristic in developing methods for estimating the operating characteristics of graded item responses (Samejima, 1977c, 1977d, 1978a, 1978b, 1978c, 1978d, 1978e, 1978f).

Bayes estimator, $\mu'_{1V}(\theta)$, of ability θ is defined by

$$(2.2) \qquad \mu'_{1V}(\theta) = \int_{-\infty}^{\infty} \theta \, P_V(\theta) \, f(\theta) \, d\theta \left[\int_{-\infty}^{\infty} P_V(\theta) \, f(\theta) \, d\theta \right]^{-1}$$

where $f(\theta)$ is the density function of θ, or the prior. This is the estimator which makes the expectation of the mean square error, such that

$$(2.3) \qquad \int_{-\infty}^{\infty} E[\theta_V^* - \theta]^2 \, f(\theta) \, d\theta \quad ,$$

where θ_V^* is any estimator of θ based upon the response pattern V, minimal (cf. Samejima, 1969).

Bayes modal estimator, $\hat{\hat{\theta}}_V$, of ability θ is the point of θ at which the function $B_V(\theta)$, which is defined by

$$(2.4) \qquad B_V(\theta) = P_V(\theta) \, f(\theta) \quad ,$$

is maximal. This estimator is similar to the maximum likelihood estimator in the sense that it maximizes a "likelihood" of a given response pattern V. Unlike the maximum likelihood estimator, however, Bayes modal estimator, as well as Bayes estimator, accompanies a certain bias which is caused by the prior, and the speed of convergence of the conditional distribution, given θ, to the unbiased normality is slower, the characteristic which will be observed and discussed in the following chapters.

Comparison of these three estimators reveals that Bayes estimator, $\mu'_{1V}(\theta)$, assumes a unique finite value under the most general condition. A sufficient condition under which a unique maximum likelihood estimate is assured for every possible response pattern has been pursued (Samejima, 1969, 1972, 1973a, 1973b, 1974), and it has been pointed out that some widely used models like the normal ogive model and the logistic model satisfy this condition, while the same is not true with the three-parameter normal ogive and logistic models (Birnbaum, 1968).

It is noted, however, that in models like the normal ogive and logistic models the maximum likelihood estimate is negative infinity for the response pattern which consists of n zeros, and that for the response pattern whose elements are the n highest item scores, m_g $(g=1,2,...,n)$, is positive infinity. A sufficient condition under which a unique Bayes modal estimate exists for every possible response pattern has also been investigated. It has been pointed out that, if, in addition to the sufficient condition for the unique maximum likelihood estimate, the first derivative of $\log f(\theta)$ is strictly decreasing in θ , a unique Bayes modal estimate exists for every possible response pattern. Unlike the maximum likelihood estimate, Bayes modal estimate is finite even for the above two extreme response patterns.

III Objective Testing and Bayesian Estimation of Ability

We can say that the purpose of objective testing is to measure an individual's ability without biases of any kinds. It is a common tendency that the graduate schools of many universities of the United States adopt the Graduate Record Examinations given by Educational Testing Service as one of the criteria in their decision of accepting or rejecting applicants, in preference to similar tests developed and used within each college. This fact can be considered as an example in which effort is taken to avoid possible biases caused by different tests and/or different norm groups, in order to measure the individual's ability objectively. It is well-known that some tests are culturally biased, and the use of such tests will result in overestimating the ability levels of individuals with some particular cultural backgrounds, and in underestimating those of individuals with some other cultural backgrounds. This second example illustrates a bias which is rooted in the contents of tests.

There is a completely different type of bias, which tends to be overlooked by psychologists and other researchers, but which affects the ability measurement just as strongly. Suppose that the content of our test is perfectly valid and unbiased. Using such a test, however, we may still result in performing a biased measurement, which is far from the purpose of objective testing, provided that we fail to choose a right method of estimating the examinees' ability levels. Thus the third type of bias is not related with the content of the test, but with the theory behind the method of estimating the examinees' ability, which we adopt in the process of analyzing our data.

We note that the maximum likelihood estimation does not, in its basis, include any information from the population to which the individual belongs, and, most importantly, there is no possibility that the resulting estimate is influenced by anything other than the examinee's performance itself. The same is not true with Bayesian estimation, however.

It has been shown (Samejima, 1969) that, using LIS-U (Indow and Samejima, 1962, 1966) and other short tests as examples, both the regression of the Bayes estimate and that of the Bayes modal estimate, on ability θ, which are given by

(3.1) $$E(\mu'_{1V}(\theta)|\theta) = \sum_V \mu'_{1V}(\theta)\, P_V(\theta)$$

and

(3.2) $$E(\hat{\hat{\theta}}_V|\theta) = \sum_V \hat{\hat{\theta}}_V\, P_V(\theta) \quad ,$$

respectively, regress toward μ, when the prior is the normal density, $n(\mu,\sigma)$. Since these two sets of results are similar to each other, in this chapter, we shall use only one estimator, i.e., Bayes modal estimator, to observe the biases.

Table 3-1 presents the discrimination parameter, a_g, and the difficulty parameter, b_g, of each of the seven binary test items of LIS-U, which follows the normal ogive model such that

$$(3.3) \qquad P_g(\theta) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{a_g(\theta-b_g)} e^{-u^2/2}\, du \ ,$$

where $P_g(\theta)$ is the operating characteristic for $x_g=1$ of the binary item g, or the item characteristic function. The item information function, $I_g(\theta)$, of item g is defined by

$$(3.4) \qquad I_g(\theta) = \sum_{x_g=0}^{m_g} I_{x_g}(\theta)\, P_{x_g}(\theta) = I_0(\theta)[1-P_g(\theta)] + I_1(\theta)P_g(\theta) \ ,$$

where $I_{x_g}(\theta)$ is the item response information function, which is given by

$$(3.5) \qquad I_{x_g}(\theta) = -\frac{\partial^2}{\partial\theta^2} \log P_{x_g}(\theta) \begin{cases} = -\dfrac{\partial^2}{\partial\theta^2} \log\,[1-P_g(\theta)] & x_g = 0 \\ = -\dfrac{\partial^2}{\partial\theta^2} \log P_g(\theta) & x_g = 1 \ . \end{cases}$$

The test information function, $I(\theta)$, can be written as the sum of the item information functions, such that

$$(3.6) \qquad I(\theta) = \sum_{g=1}^{n} I_g(\theta) \ .$$

TABLE 3-1

Item Discrimination Parameter, a_g , and Item Difficulty Parameter, b_g , of Each of the Seven Items of LIS-U .

Item	a_g	b_g
1	1.031	-0.860
2	1.695	-0.520
3	1.020	-0.220
4	0.800	-0.030
5	1.111	0.190
6	1.389	0.470
7	1.370	0.760

Figure 3-1 presents the test information function of LIS-U , and its square root, which is considered as the reciprocal of the standard error of estimation defined as a function of ability θ .

The regression of the Bayes modal estimator, $\hat{\theta}_V$, on ability θ , which is given in (3.2), was obtained by using each of the four different priors, n(0.0,1.0) , n(-1.0,1.0) , n(1.0,1.0) and n(0.0,0.5) . These four regressions are shown in Figure 3-2. For convenience, hereafter, we shall call these four cases Cases 1, 2, 3 and 4 .

We can see in Figure 3-2 that these four conditional means of $\hat{\hat{\theta}}$ are substantially different from one another, in spite of the fact that they are all estimates of θ obtained through the same test, LIS-U . We note, in addition, that none of these four regressions is close to the straight line, which is drawn by a solid line in Figure 3-2 indicating the unbiasedness of estimation, and the discrepancies are large for values of θ which are far from the mean of each prior. Discrepancies among the four conditional means are great even at $\theta = 0$, where the test information function, $I(\theta)$, of LIS-U assumes as high a value as 5.55546 ; the fact which indicates strong biases of estimation, i.e., the expectation of ability is 0.00299, -0.16252, 0.16915 or 0.00126 , depending upon the prior to which examinees of ability 0.0 are assigned to. Thus the examinees who belong to Case 2 are

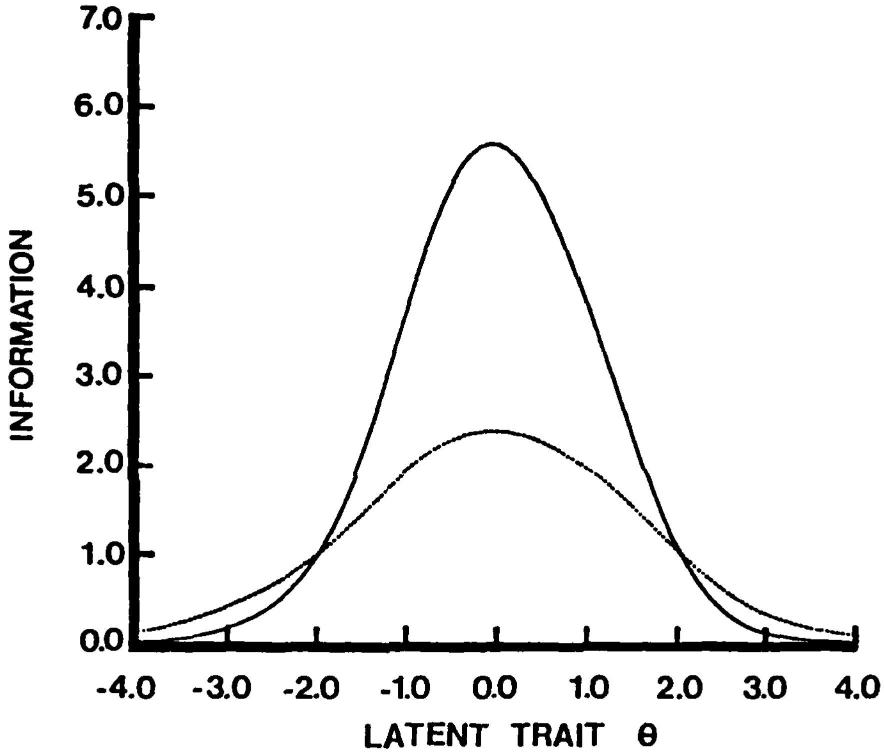

FIGURE 3-1

Test Information Function (Solid Line) and Its Square Root (Dotted Line) of LIS-U .

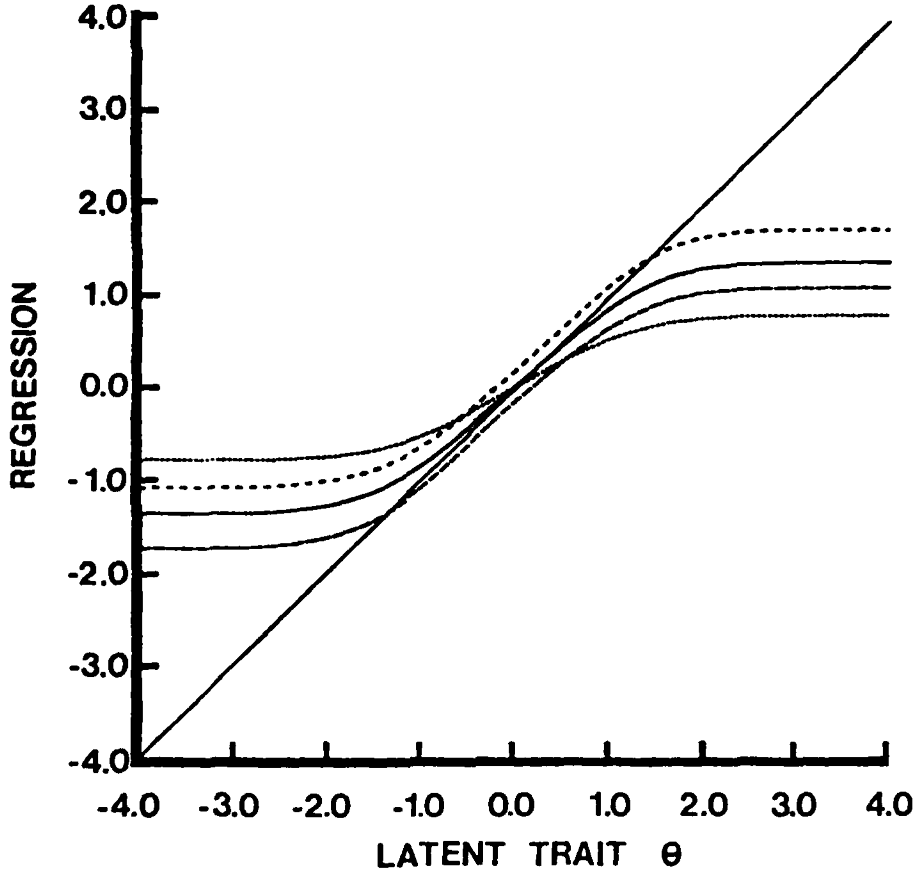

FIGURE 3-2

Four Regressions of the Bayes Modal Estimate on Ability Based on LIS-U, with the Priors, n(0.0,1.0) (Solid Line), n(-1.0,1.0) (Broken Line), n(1.0,1.0) (Dashed Line), and n(0.0,0.5) (Dotted Line), Respectively.

severely and unqualifiedly handicapped while those who belong to Case 3 are allowed to enjoy the advantage, <u>regardless of the fact that they are individuals whose ability levels are exactly the same</u>.

These discrepancies in regression are enhanced if we shift the true ability level from 0.0 to 0.5 , at which the test information function assumes 4.93877 , slightly less than 5.55546 at $\theta = 0.0$. The expected values of the Bayes modal estimate are 0.44410, 0.26827, 0.63169 and 0.28436 for Cases 1, 2, 3 and 4, respectively, providing us with the range of 0.36342 . Greater discrepancies are observed, however, at levels of θ where the amount of test information is much smaller. At $\theta = 3.0$ where $I(\theta) = 0.08634$, for example, the expected Bayes modal estimates are 1.34995, 1.07489, 1.70326 and 0.77629 , respectively, none of which is close to the true ability level, 3.0 ; at $\theta = -2.0$ where $I(\theta) = 0.96397$, the expected estimates are -1.26590, -1.61087, -1.00075 and -0.73634 . We can see that at these ability levels Bayes modal estimate is at the mercy of a given prior; the fact which is against the principle of objective testing.

We note that the expected Bayes modal estimates are 0.00299 and 0.00126 at $\theta = 0.0$ for Cases 1 and 4, respectively, both of which are very close to the true ability level, 0.0 , whereas in Cases 2 and 3 the expected estimates are -0.16252 and 0.16915 , which are farther from 0.0 in the two directions. These biases for Cases 1 and 4 come from the priors, i.e., n(0.0,1.0) and n(0.0,0.5) respectively, showing regressions toward the means of the separate priors. Similar tendencies are observed at $\theta = -1.0$ and $\theta = 1.0$, where the means of the priors for Cases 2 and 3 are located, respectively; the expected Bayes modal estimates are -0.85033, -1.09301, -0.64375 and -0.52486 at $\theta = -1.0$, and 0.84548, 0.63805, 1.08443 and 0.51965 at $\theta = 1.0$, for Cases 1, 2, 3, and 4.

It is evident from the above observations that, _even if the test itself is perfectly objective in content, the use of Bayes modal estimator of ability will destroy the objectivity of testing_, providing the examinees with unqualified advantages or disadvantages, depending upon the relative positions of their ability levels and the prior to which they are assigned.

As was mentioned earlier, the maximum likelihood estimator is asymptotically unbiased, the characteristic which suits the principle of objective testing, although for short tests the approximation may not be very good. It will be worthwhile, therefore, to investigate the destruction of objectivity by the Bayes modal estimator in comparison with the behavior of the maximum likelihood estimator.

Figure 3-3 presents four functions, i.e., the standard normal density function, n(0,1) (solid line), and three approximations to n(0,1). Each of these three approximations is the product of two functions, $P_h(\theta)$ and $[1-P_j(\theta)]$, which are given by the normal ogive functions such that

$$(3.7) \qquad P_h(\theta) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{a_h(\theta-b_h)} e^{-u^2/2}\, du$$

and

$$(3.8) \qquad P_j(\theta) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{a_j(\theta-b_j)} e^{-u^2/2}\, du \quad ,$$

where $a_h = a_j$ and $b_h = -b_j$. These two parameters, a_h and b_h , are 0.94810 and -0.35454 for the function drawn by a dotted line in Figure 3-3, 0.94980 and -0.35391 for the one drawn by a broken or long, dashed line, and 0.95259 and -0.35287 for the one drawn by a short, dashed line, respectively. These three approximations are obtained by setting the product of the two functions equal to the standard normal density function at $\theta = 0.3$, $\theta = 0.6$ and $\theta = 0.9$, respectively, in addition to $\theta = 0.0$. We notice that these four curves, including $n(0,1)$, in Figure 3-3 are practically indistinguishable.

We notice that the formulas in (3.7) and (3.8) are identical with the item characteristic function in the normal ogive model on the dichotomous response level, which is shown as (3.3). This implies that the prior, $n(0,1)$, is practically the same

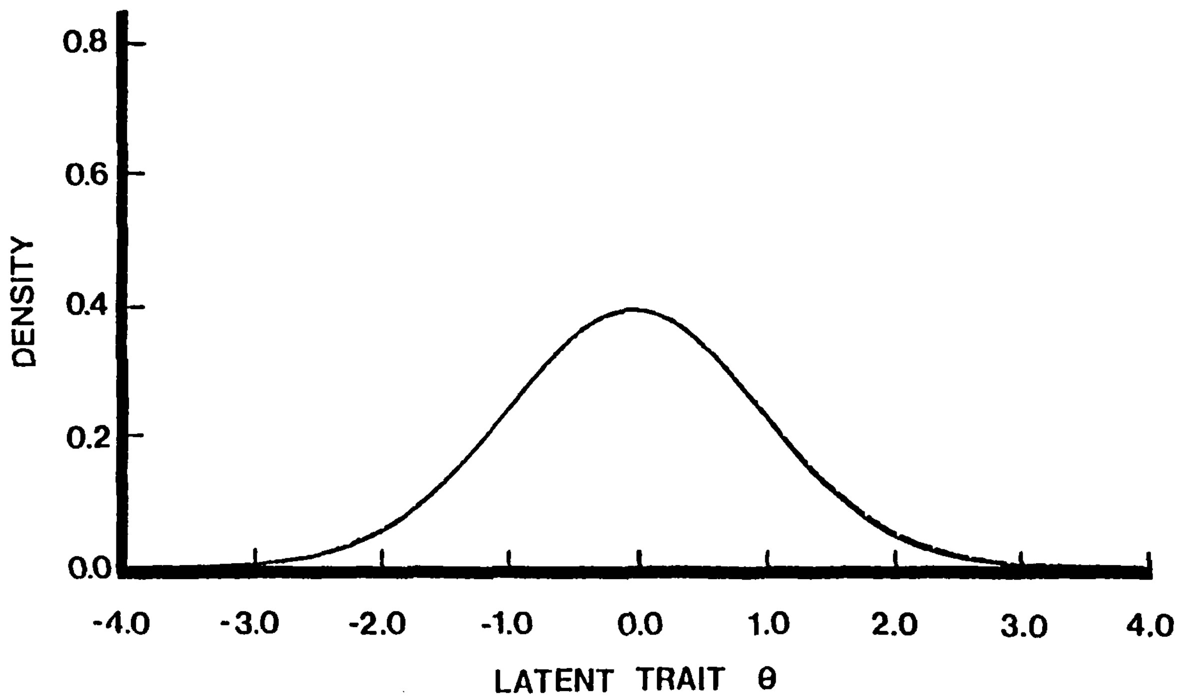

FIGURE 3-3

Comparison of Three Approximations with the Normal Density Function, n(0,1) (Solid Line). These Approximations Are the Products of a Normal Ogive Function and Another Subtracted from Unity, Which Equal n(0,1) at $\theta = 0.3$ (Dotted Line), $\theta = 0.6$ (Broken Line) and $\theta = 0.9$ (Dashed Line), Respectively.

as the product of the two operating characteristics of the hypothetical binary items, h and j, for the response pattern, (1,0) . The Bayes modal estimator with the prior $n(0,1)$ can be considered, therefore, as the maximum likelihood estimator, obtained from the response pattern on LIS-U <u>plus additional two responses</u>, 1 <u>and</u> 0 , to the hypothetical binary items, h and j . Note that these two additional item responses are always 1 and 0 , regardless of the true ability level.

Let V^* denote any response pattern on the two hypothetical test items, h and j . Since both are binary items, there are only four possible response patterns V^* , i.e., (0,0), (0,1), (1,0) and (1,1) . The operating characteristic, $P_{V^*}(\theta)$, of the response pattern V^* is given by

$$(3.9) \qquad P_{V^*}(\theta) \begin{cases} = [1-P_h(\theta)][1-P_j(\theta)] & \text{for } V^{*\prime}=(0,0) \\ = [1-P_h(\theta)]\ P_j(\theta) & \text{for } V^{*\prime}=(0,1) \\ = P_h(\theta)\ [1-P_j(\theta)] & \text{for } V^{*\prime}=(1,0) \\ = P_h(\theta)\ P_j(\theta) & \text{for } V^{*\prime}=(1,1) \end{cases},$$

where $P_h(\theta)$ and $P_j(\theta)$ are the item characteristic functions of the hypothetical binary items, h and j , which are given by (3.7) and (3.8), respectively. Figure 3-4 presents the operating characteristics of the four response patterns with $a_h = a_j = 0.95$ and $b_h = -b_j = -0.35$.

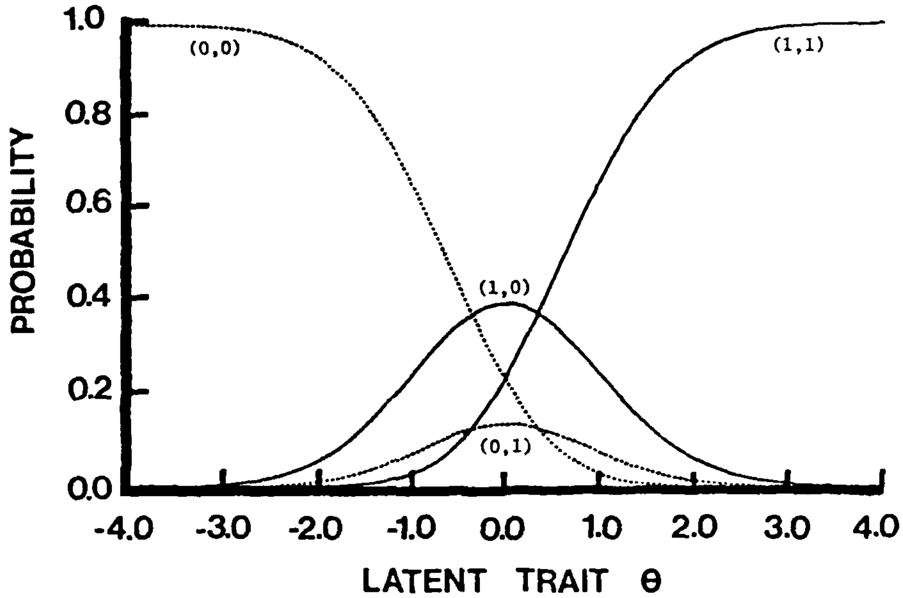

FIGURE 3-4

Operating Characteristics, $P_{V*}(\theta)$, of Four Possible Response Patterns, (0,0) , (0,1) , (1,0) and (1,1) , of Hypothetical Binary Items, h and j .

As we can see in Figure 3-4, although the probability is highest at $\theta = 0.0$ for $V^{*\prime} = (1,0)$ compared with those for the other three, i.e., 0.397 against 0.233 for $V^{*\prime} = (0,0)$, 0.137 for $V^{*\prime} = (0,1)$ and 0.233 for $V^{*\prime} = (1,1)$, and this tendency holds in the vicinity of this ability level, at $\theta = -0.4$ and $\theta = 0.4$ it is already exceeded by the probabilities for $V^{*\prime} = (0,0)$ and $V^{*\prime} = (1,1)$, respectively, i.e., 0.367 against 0.395 in each case. If we shift the ability level from 0.0 to ± 1.0, ± 1.5, ± 2.0, ± 2.5, ± 3.0 and ± 3.5, the probability for $V^{*\prime} = (1,0)$ decreases rapidly relative to either the probability for $V^{*\prime} = (1,1)$ or the one for $V^{*\prime} = (0,0)$, i.e., 0.242 vs. 0.659, 0.132 vs. 0.829, 0.058 vs. 0.929, 0.020 vs. 0.976, 0.006 vs. 0.993 and 0.001 vs. 0.998, respectively. In other words, at $\theta = 2.0$, for example, chances are only 58 times out of 1,000 that the examinee of this ability level obtains (1,0) for $V^{*\prime}$, in comparison with 929 times out of 1,000 for $V^{*\prime} = (1,1)$. As far as we use Bayes modal estimator, however, it is treated as if chances were 1,000 times out of 1,000 for the examinee of this ability level to obtain (1,0) for $V^{*\prime}$! It is no wonder that the conditional expectation of the Bayes modal estimate, given θ, regresses toward the center, which indicates the tendency that examinees of lower ability levels obtain higher values of the Bayes modal estimate and those of higher levels obtain lower values of the Bayes modal estimate. The effect must be especially strong when the number of items in the test is relatively small. Note that this tendency is relative to the prior, to which the individual is assigned. In other words, the hypothetical test items h and j, whose item characteristic functions approximate the prior, differ from one population to another, the fact which explains the relative positions of the four regressions in Figure 3-2. This additional response pattern, (1,0), for the hypothetical test items h and j creates nothing but biases which contradict the principle of objective testing.

Although, in the above examples, normal density functions were solely used for priors, we can see that the same logic can be applied for priors of different shapes. The resultant bias caused by Bayesian estimation depends upon the particular shape of the prior, and the set of hypothetical items and the specific response pattern, whose operating characteristic approximates the prior.

IV Bayes Modal Estimate When the Amount of Test Information Is Large

In the preceding chapter, we observed the bias caused by the Bayesian estimation using Bayes modal estimator with a relatively short test, LIS-U. Since most priors can be approximated by the product of the operating characteristics of a relatively small number of hypothetical items, it is expected that the effect of a prior will be less dominating in the resultant estimation of ability if the test is longer and more informative, i.e., if the test information function assumes high values for the entire range of ability of our interest. In this chapter, therefore, we shall observe the effect of priors in Bayesian estimation using two hypothetical tests, Test A and Test B, each of which provides us with an approximately constant amount of test information, 21.6 , for the interval of θ , [-3.0, 3.0] (cf. Samejima, 1977c). Test A consists of thirty-five graded test items with $m_g=2$ for each item, while Test B consists of twenty items with $m_g=3$ for every item. All of these graded test items follow the normal ogive model on the graded response level, whose operating characteristic, $P_{x_g}(\theta)$, of the item score x_g $(=0,1,\ldots,m_g)$ is given by

$$(4.1) \qquad P_{x_g}(\theta) = \frac{1}{\sqrt{2\pi}} \int_{a_g(\theta-b_{x_g+1})}^{a_g(\theta-b_{x_g})} e^{-u^2/2}\, du ,$$

where

$$(4.2) \qquad -\infty = b_0 < b_1 < \ldots\ldots < b_{m_g} < b_{m_g+1} = \infty ,$$

and

(4.3) $a_g > 0$.

Table 4-1 presents the discrimination parameter, a_g , and the two difficulty parameters, b_{x_g} for $x_g = 1, 2$, of each of the thirty-five items of Test A. Table 4-2 also presents the discrimination parameter, a_g , and the three difficulty parameters, b_{x_g} for $x_g = 1, 2, 3$, of each of the twenty items of Test B.

For each of the one hundred hypothetical examinees, whose ability θ distributes approximately normally (cf. Samejima, 1977c), both the maximum likelihood estimate and Bayes modal estimate were obtained upon a response pattern, which was calibrated by the Monte Carlo method, for each of Tests A and B. Figure 4-1 presents these two estimates plotted against the true ability θ for each of the one hundred hypothetical examinees. We can see in these two graphs of Figure 4-1 that Bayes modal estimate, which is represented by solid triangles, tends to regress toward the center, in comparison with the maximum likelihood estimate, which is drawn by crosses, for both Tests A and B, although the tendency is less conspicuous than in the case of LIS-U. The sample linear regression of any estimator θ^* on ability θ is given by $\alpha_0 + \alpha_1 \theta$, where

(4.4) $\alpha_0 = M_{\theta*} - (s_{\theta*}/s_\theta) \text{ Corr.}(\theta,\theta^*) M_\theta$

and

TABLE 4-1

Item Discrimination Parameters, a_g , and the Two Item Difficulty Parameters, b_1 and b_2 , of Each of the Thirty-Five Graded Items of Test A.

Item g	a_g	b_1	b_2
1	1.8	-4.75	-3.75
2	1.9	-4.50	-3.50
3	2.0	-4.25	-3.25
4	1.5	-4.00	-3.00
5	1.6	-3.75	-2.75
6	1.4	-3.50	-2.50
7	1.9	-3.00	-2.00
8	1.8	-3.00	-2.00
9	1.6	-2.75	-1.75
10	2.0	-2.50	-1.50
11	1.5	-2.25	-1.25
12	1.7	-2.00	-1.00
13	1.5	-1.75	-0.75
14	1.4	-1.50	-0.50
15	2.0	-1.25	-0.25
16	1.6	-1.00	0.00
17	1.8	-0.75	0.25
18	1.7	-0.50	0.50
19	1.9	-0.25	0.75
20	1.7	0.00	1.00
21	1.5	0.25	1.25
22	1.8	0.50	1.50
23	1.4	0.75	1.75
24	1.9	1.00	2.00
25	2.0	1.25	2.25
26	1.6	1.50	2.50
27	1.7	1.75	2.75
28	1.4	2.00	3.00
29	1.9	2.25	3.25
30	1.6	2.50	3.50
31	1.5	2.75	3.75
32	1.7	3.00	4.00
33	1.8	3.25	4.25
34	2.0	3.50	4.50
35	1.4	3.75	4.75

TABLE 4-2

Item Discrimination Parameters, a_g , and the Three Item Difficulty Parameters, b_1 , b_2 and b_3 , of Each of the Twenty Graded Items of Test B.

Item g	a_g	b_1	b_2	b_3
1	1.0	-5.5	-4.5	-3.5
2	1.3	-5.0	-4.0	-3.0
3	2.2	-4.5	-3.5	-2.5
4	2.2	-4.1	-3.1	-2.1
5	2.5	-3.7	-2.7	-1.7
6	2.8	-3.0	-2.0	-1.0
7	1.9	-2.6	-1.6	-0.6
8	1.6	-2.0	-1.0	0.0
9	1.3	-1.5	-0.5	0.5
10	1.6	-1.1	-0.1	0.9
11	2.5	-1.0	0.0	1.0
12	2.8	-0.7	0.3	1.3
13	1.9	-0.2	0.8	1.8
14	2.2	1.2	2.2	3.2
15	1.6	1.4	2.4	3.4
16	2.5	1.5	2.5	3.5
17	1.9	1.7	2.7	3.7
18	2.2	2.1	3.1	4.1
19	2.8	2.8	3.8	4.8
20	1.0	3.5	4.5	5.5

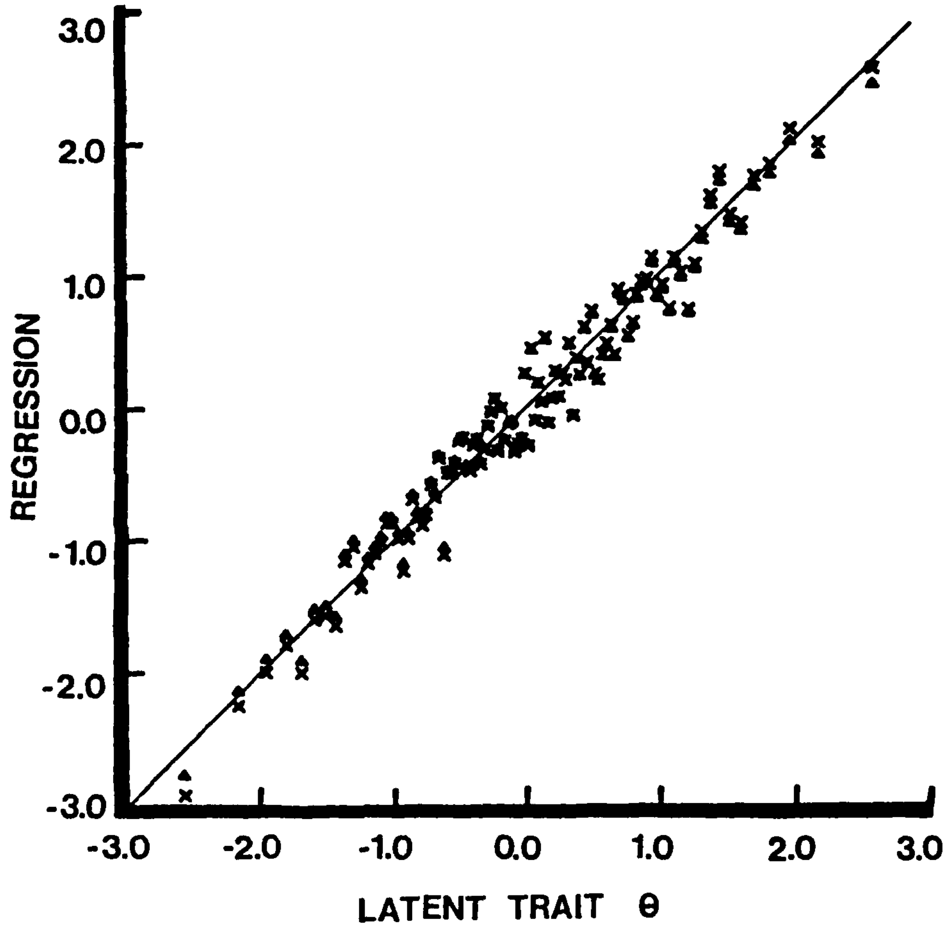

FIGURE 4-1

Maximum Likelihood Estimates (Crosses) and Bayes Modal Estimates (Triangles) of Ability for One Hundred Hypothetical Examinees Whose Ability Distributes Approximately N(0,1) , Obtained Through Test A .

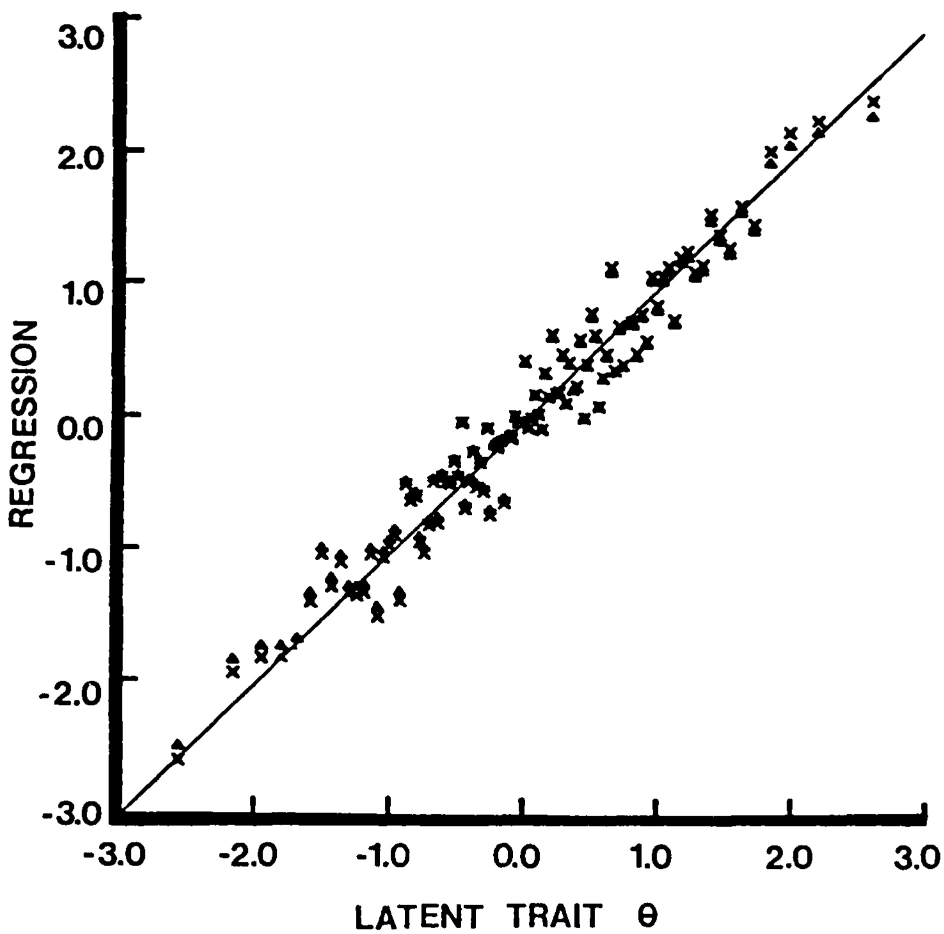

FIGURE 4-1 (Continued): Those Obtained through Test B ·

(4.5) $\alpha_1 = (s_{\theta*}/s_\theta)$ Corr.(θ,θ^*) ,

with M and s representing the sample mean and the sample standard deviation, respectively. Replacing θ^* by the maximum likelihood estimate $\hat{\theta}$, the two coefficients, α_0 and α_1 , of the sample linear regression were calculated for each of Tests A and B, and are presented in Table 4-3. We can see from this result that the sample linear regression for Test A is almost identical with the straight line drawn in Figure 4-1, indicating the unbiased estimation, since 0.9933 is very close to unity and -0.0088 to zero. For Test B, the sample linear regression is flatter than the line of forty-five degrees, showing regression toward the center, although the degree of regression is very small. The corresponding set of coefficients were calculated for the Bayes modal estimate, $\hat{\hat{\theta}}$, for both Tests A and B, and are shown in the same table. We can see that for both tests Bayes modal estimate tends to regress toward the center more strongly. This tendency is far less than it is for a short test like LIS-U, however, the fact which is anticipated from the difference in the amounts of test information for this range of θ .

As another example, Bayes modal estimate was obtained for each of five hundred hypothetical examinees, whose ability levels differ from -2.475 to 2.475 with the step of 0.05 , with five examinees sharing each ability level. The estimation was made upon the response pattern of Test A, which was calibrated by the

TABLE 4-3

Coefficients of the Linear Regressions of the Maximum Likelihood Estimate (MLE) and the Bayes Modal Estimate (BME) on Ability θ , for One Hundred Observations on Each of Tests A and B.

		α_0	α_1
Test A	MLE	-0.0088	0.9933
	BME	-0.0081	0.9476
Test B	MLE	0.0025	0.9617
	BME	0.0027	0.9201

Monte Carlo method, using each of the four priors, $n(0.0,1.0)$, $n(-1.0,1.0)$, $n(1.0,1.0)$ and $n(0.0,0.5)$.

Figure 4-2 presents the mean of the five Bayes modal estimates thus obtained for each of the one hundred levels of θ, which is plotted by a dot, together with the corresponding mean of the five maximum likelihood estimates, which is represented by a cross, for each of the four priors. Examination of each of these four graphs reveals the anticipated bias of the Bayesian estimation, i.e., the tendency to regress toward 0.0, -1.0, 1.0 and 0.0, respectively, although it is even less conspicuous than in the preceding example, except for one case in which the prior is $n(0.0,0.5)$. The five sets of coefficients of the sample linear regressions of the five hundred estimates on ability were calculated in the same manner as in the preceding example, and are shown in Table 4-4. Again the sample linear regression of the maximum likelihood estimate is practically identical with the straight line of forty-five degrees, with the two coefficients, $\alpha_0 = -0.0058$ and $\alpha_1 = 1.0047$, being so close to zero and unity, respectively, while the other sets of coefficients for the four sets of Bayes modal estimates indicate flatter lines, suggesting separate and anticipated regressions.

From these results, it is obvious that even with an informative test having a large amount of test information, like 21.6, for the entire range of ability θ of our interest the

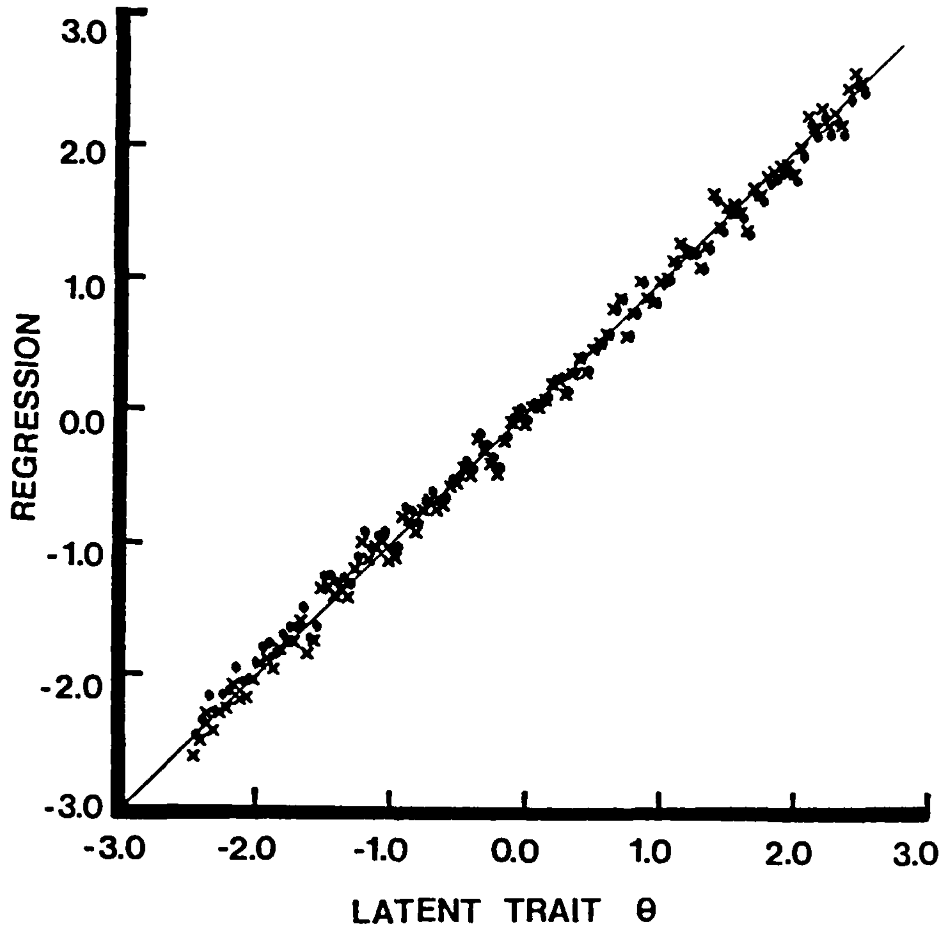

FIGURE 4-2

Conditional Mean of the Five Maximum Likelihood Estimates (Crosses), and that of the Five Bayes Modal Estimates (Dots), of Ability for Each of the One Hundred Levels of Ability for Test A. The Prior for the Bayes Modal Estimates Is n(0,1) .

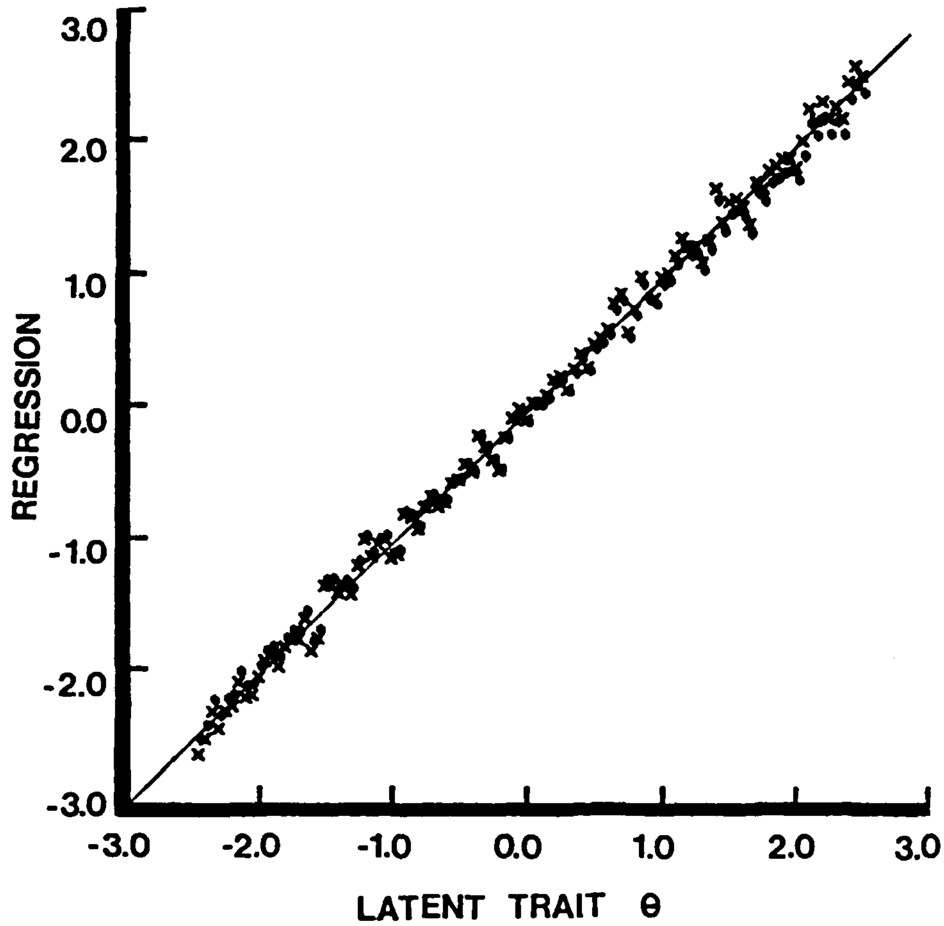

FIGURE 4-2 (Continued): The Prior is n(-1,1) .

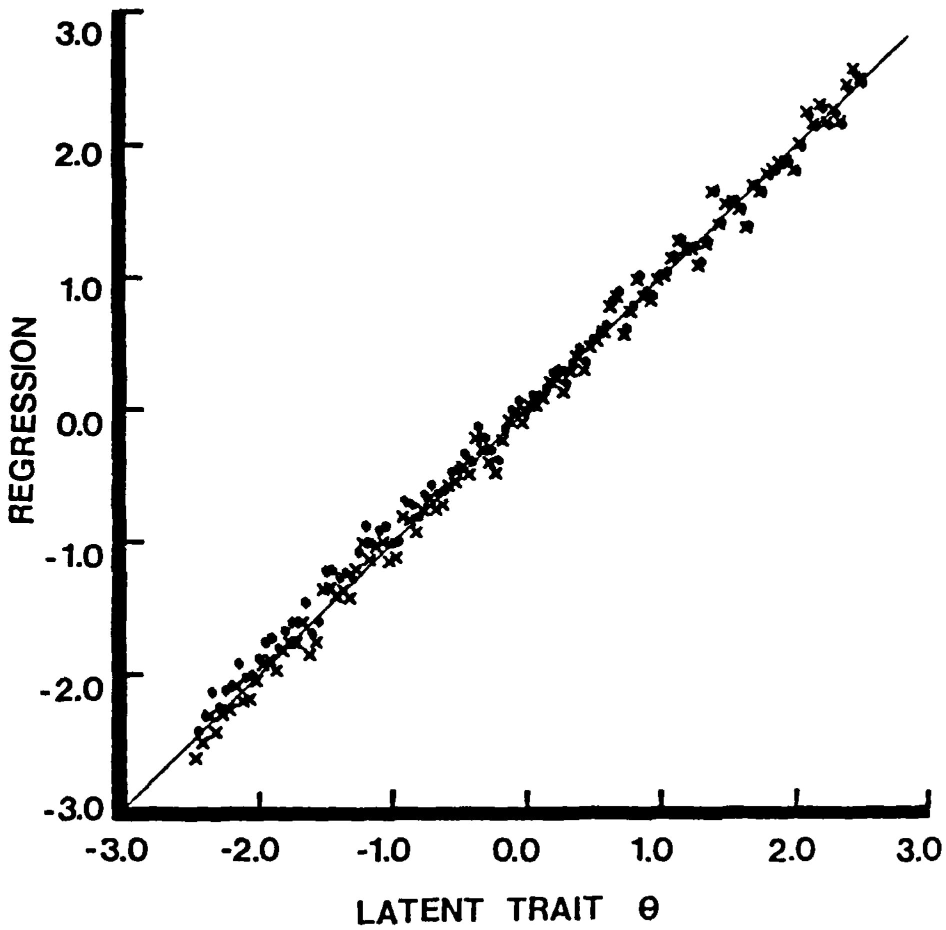

FIGURE 4-2 (Continued): The Prior is n(1,1) .

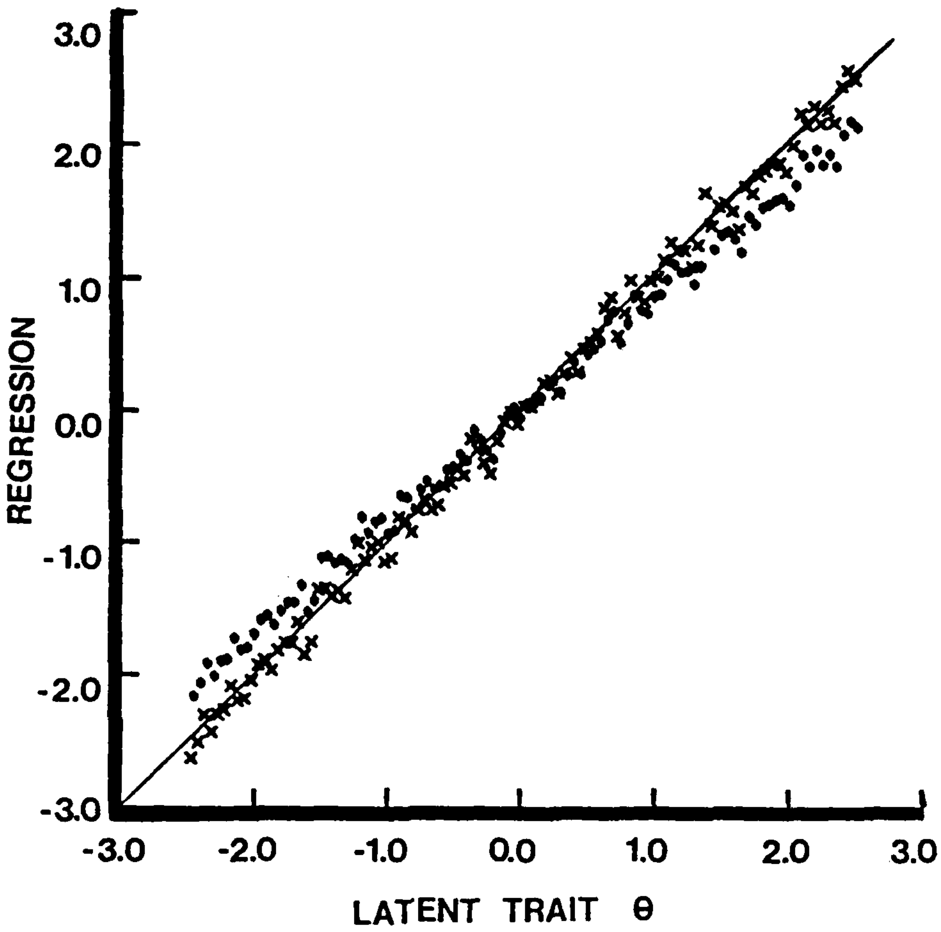

FIGURE 4-2 (Continued): The Prior is n(0.0,0.5) .

TABLE 4-4

Coefficients of the Linear Regression of Each of the Five Estimators on Ability θ , for Five Hundred Observations on Test A. The Five Estimators are the Maximum Likelihood Estimator (MLE) and the Four Bayes Modal Estimators with Different Priors.

	α_0	α_1
MLE	-0.0058	1.0047
BME		
n(0.0,1.0)	-0.0053	0.9591
n(-1.0,1.0)	-0.0500	0.9592
n(1.0,1.0)	0.0395	0.9590
n(0.0,0.5)	-0.0039	0.8456

effect of a prior in Bayes modal estimation appears in the form of bias, which was observed with a shorter test like LIS-U whose maximum amount of test information is 5.55546 . We can see that, as we increase the amount of test information, Bayes modal estimate approaches the maximum likelihood estimate. This implies that Bayes modal estimate, too, has the asymptotic unbiasedness, as the maximum likelihood estimate does. The convergence to the unbiasedness is slower for the Bayes modal estimate, however, since Bayes modal estimate must "shake off" the effect of the prior in the process of approaching the unbiasedness. We can say that the prior is nothing but an obstacle whose effect should be gotten rid of in order to approach the unbiasedness of estimation, which is essential for objective testing.

V Effect of the Prior in Tailored Testing

We shall observe here how the prior affects the resultant ability estimation in tailored testing, where a single item is selected from an item pool and presented to an individual examinee, sequentially. A tailored testing situation is simulated with an hypothetical item pool, in which there are nine binary item groups, each of which consists of a large number of equivalent, binary test items following the normal ogive model, which is given by (3.3) and whose item discrimination parameter, a_g , and item difficulty parameter, b_g , for the item group g are shown in Table 5-1. We assume eleven hypothetical examinees, whose ability levels are -2.25, -1.75, -1.25, -0.75, -0.25, 0.00, 0.50, 1.00, 1.50, 2.00 and 2.50, respectively. We also assume four different situations, in one of which the maximum likelihood estimation is applied for the ability estimation, and in the other three Bayes modal estimation is used, with three different priors, $n(0.0,1.0)$, $n(0.0,0.8)$ and $n(0.0,0.5)$, respectively. In the first situation of maximum likelihood estimation, an item from group 5 is always chosen as the first item to present to an examinee, and, depending upon the examinee's response to this item, the second item is chosen either from group 1 or group 9. That is to say, if the examinee's response to the first item is correct, then the second item is chosen from group 9, i.e., the most

TABLE 5-1

Item Discrimination Parameter, a_g , and Item Difficulty Parameter, b_g , of Each of the Nine Groups of Binary Test Items Used as the Item Pool in the Simulated Tailored Testing.

Item Group	a_g	b_g
1	1.20	-2.00
2	1.60	-1.50
3	2.00	-1.00
4	1.40	-0.50
5	1.80	0.00
6	1.30	0.50
7	1.70	1.00
8	1.90	1.50
9	1.50	2.00

difficult item group, and, if it is incorrect, then the second item is chosen from group 1, the easiest item group. The examinee will stay with the same item group for the following items, until he fails in answering an item correctly if it is group 1, and until he succeeds in answering an item incorrectly if it is group 9. Thereafter, since every current likelihood function has a local maximum, an item from the item group whose item information function, $I_g(\theta)$, which is defined by (3.4), is the greatest at the value of current maximum likelihood estimate is chosen and presented next, and this will go on until the amount of test information at the current maximum likelihood estimate reaches or exceeds a certain criterion. All the responses of the hypothetical examinees are calibrated by the Monte Carlo method.

In Bayesian estimation, the first estimate is the modal point of the prior. The second item is an item chosen from the item group whose item information function, $I_g(\theta)$, is the greatest at the modal point of the prior, and the third item is from the item group whose item information function is the greatest at the current Bayes modal estimate, and so forth, and the presentation of a new item is terminated when the amount of test information at the current estimate of the examinee's ability has reached the same criterion used in the maximum likelihood estimation.

Figures 5-1 through 5-3 present the result of these four simulated tailored testing for each of the eleven hypothetical examinees. In each of these three figures, the sequential result of the maximum likelihood estimation is presented by solid triangles, and that of one of the three Bayesian estimations is shown by hollow circles. In each Bayesian estimation, the first circle is located at the modal point of the prior, so the actual number of test items used in the simulated tailored testing is one less than the number of circles. The number of test items which are presented to each examinee in each situation is shown in parentheses in Table 5-2, following the eventual ability estimate. The amount of test information used as the criterion for terminating the presentation of a new item in this simulated tailored testing is 20.0.

As was the case with the previous examples, the effect of a prior appears in the form of underestimating the ability levels of examinees which are much higher than the mean of the prior, and of overestimating those which are much lower, in all three cases of the Bayesian estimation, with some exceptions at $\theta = -2.25$. Note that, in comparison with these results, errors of measurement in the maximum likelihood estimation are more randomly distributed in both directions. We notice, moreover, that even in the two exceptions at $\theta = -2.25$, it

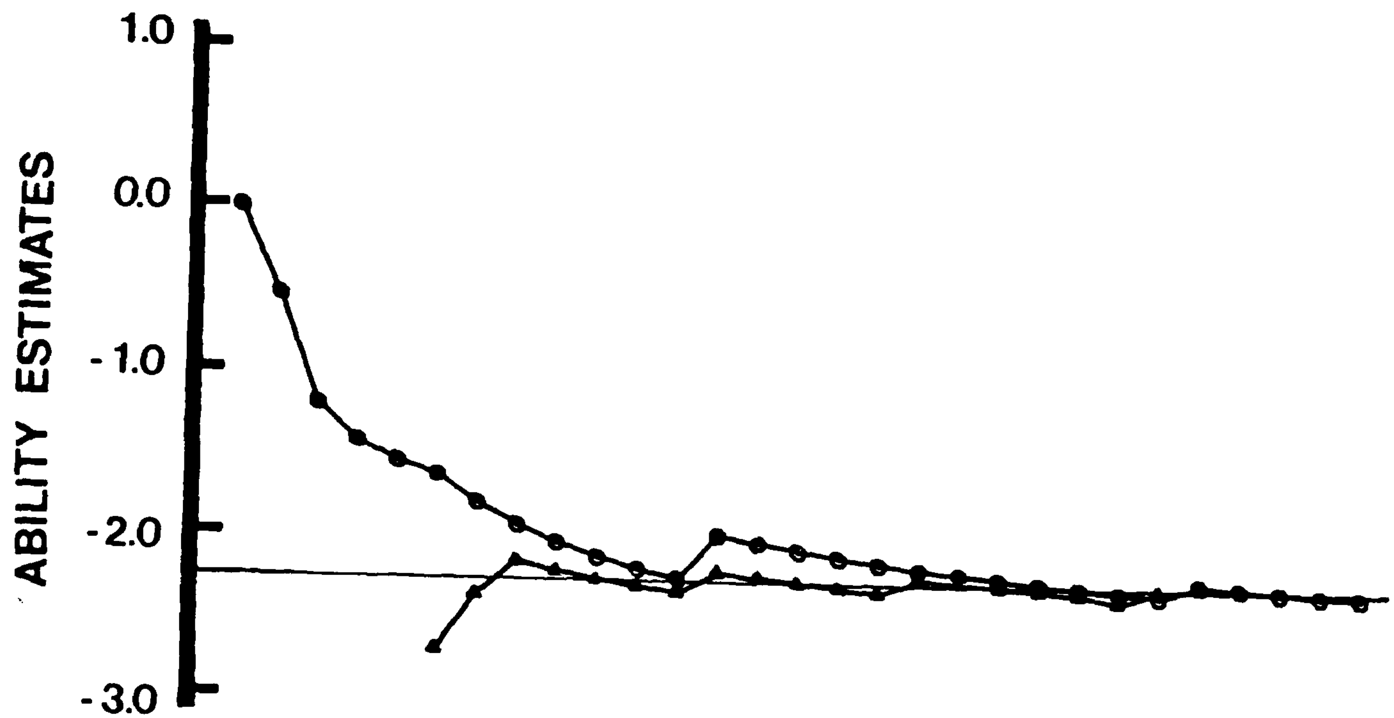

FIGURE 5-1

Successive Maximum Likelihood Estimates (Triangles) and Bayes Modal Estimates (Circles) in the Simulated Tailored Testing with n(0,1) as the Prior, for a Hypothetical Examinee Whose Ability Level is -2.25 .

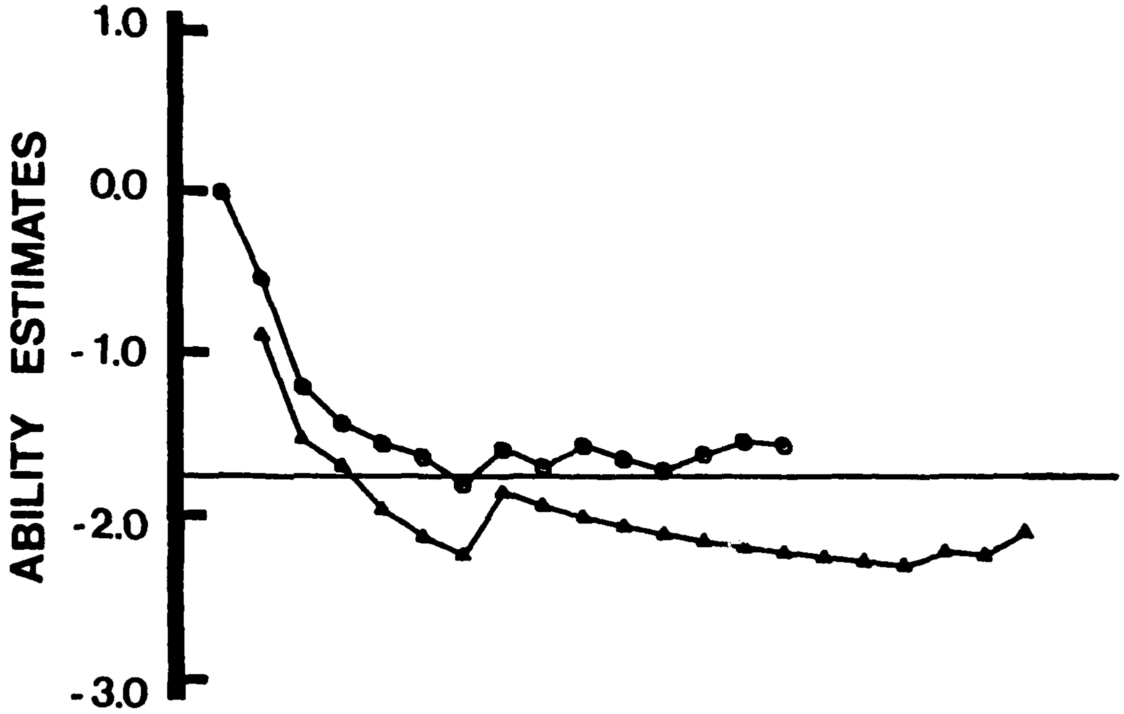

FIGURE 5-1 (Continued): The Prior is n(0,1) , θ = -1.75 .

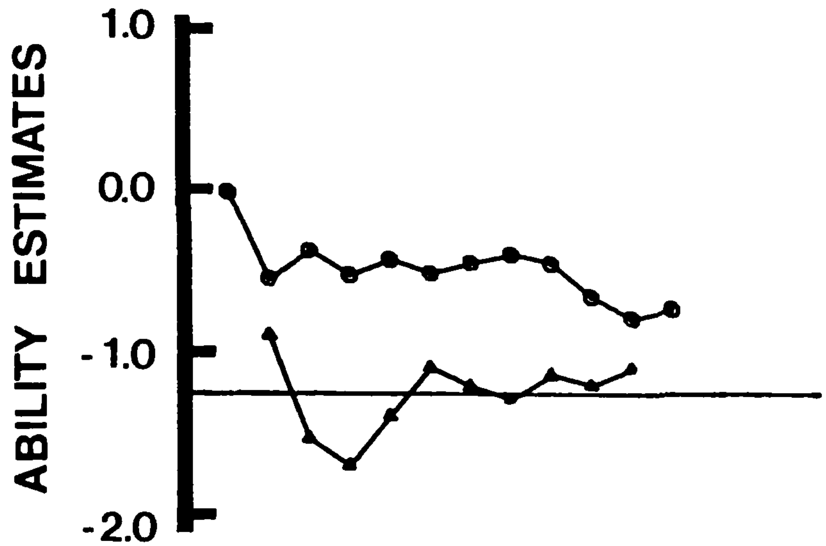

FIGURE 5-1 (Continued): The Prior is $n(0,1)$, $\theta = -1.25$.

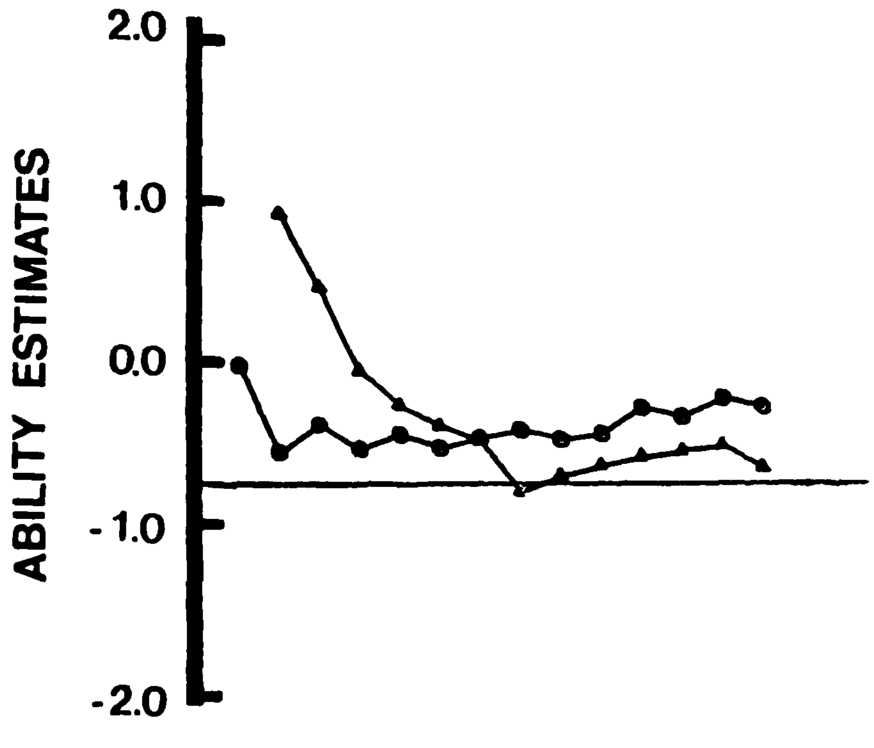

FIGURE 5-1 (Continued): The Prior is n(0,1) , $\theta = -0.75$.

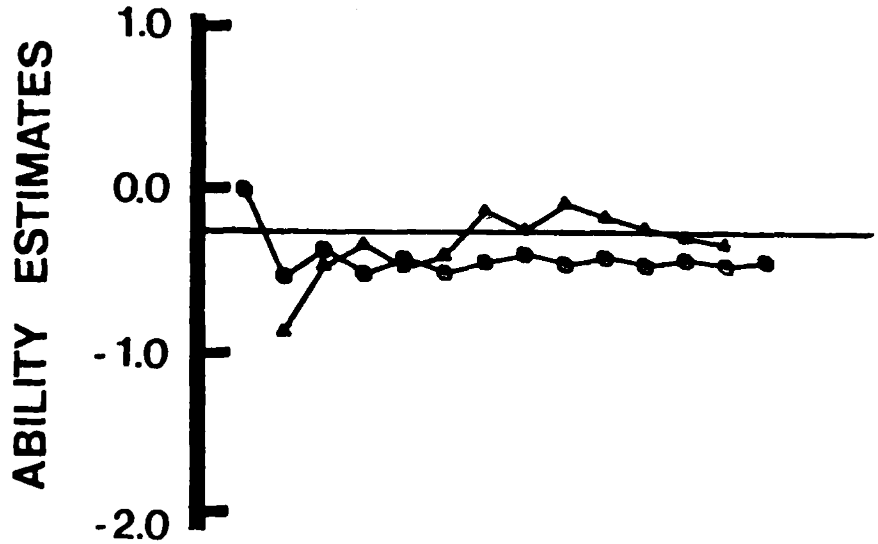

FIGURE 5-1 (Continued): The Prior is $n(0,1)$, $\theta = -0.25$.

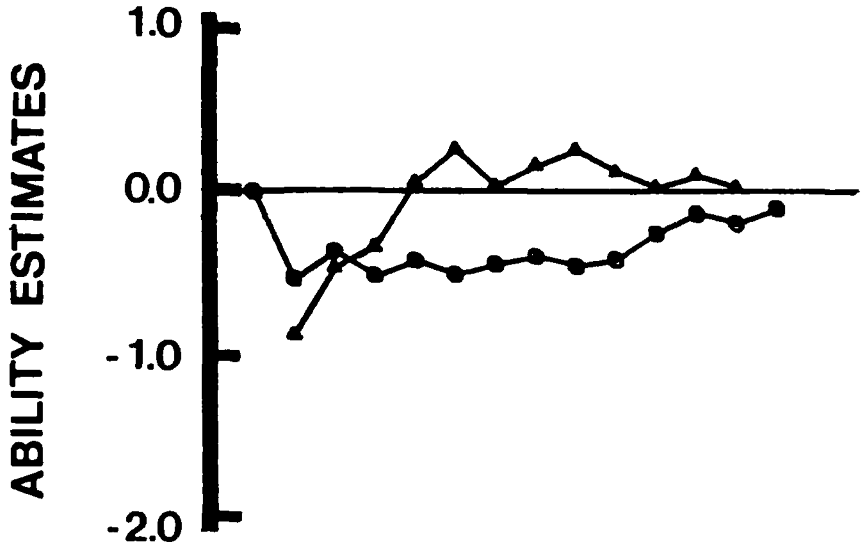

FIGURE 5-1 (Continued): The Prior is n(0,1) , θ = 0.00 .

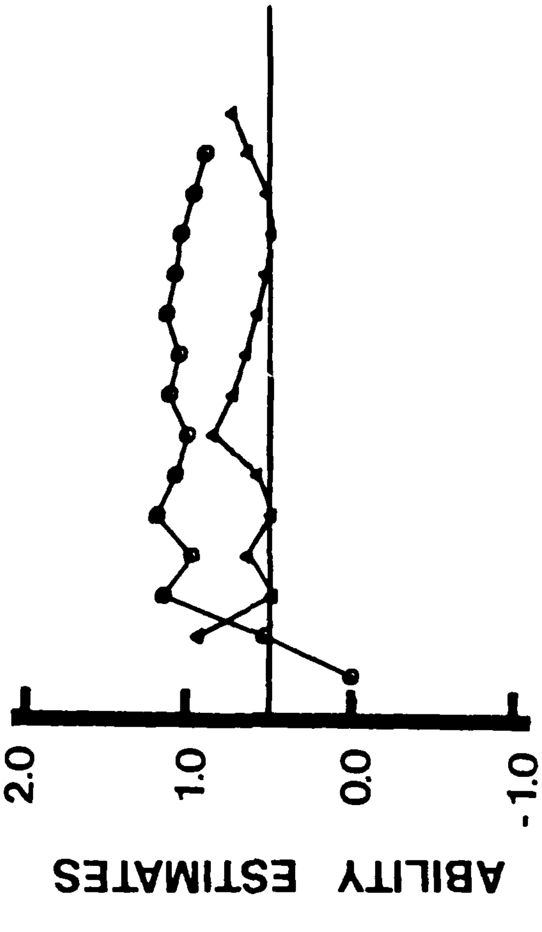

FIGURE 5-1 (Continued): The Prior is n(0,1) , θ = 0.50 .

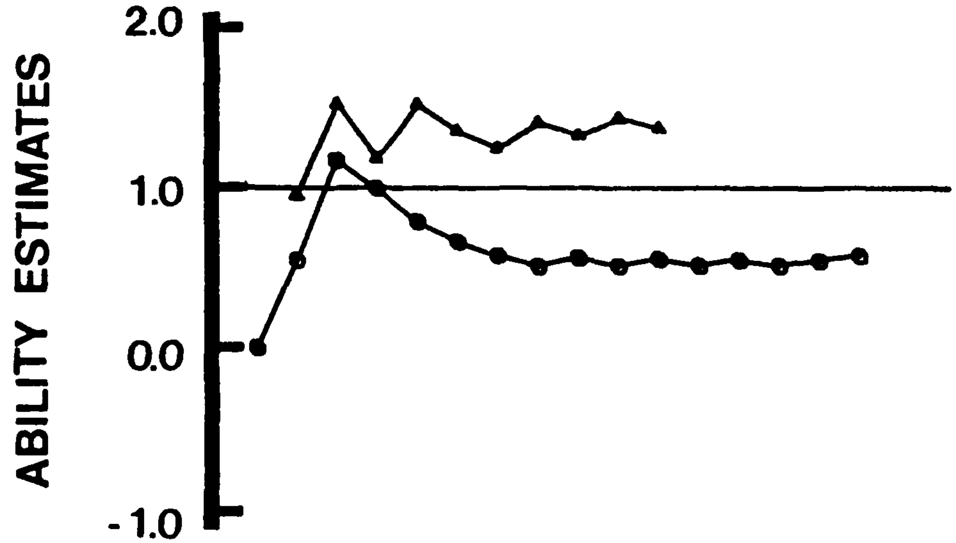

FIGURE 5-1 (Continued): The Prior is n(0,1) , θ = 1.00 .

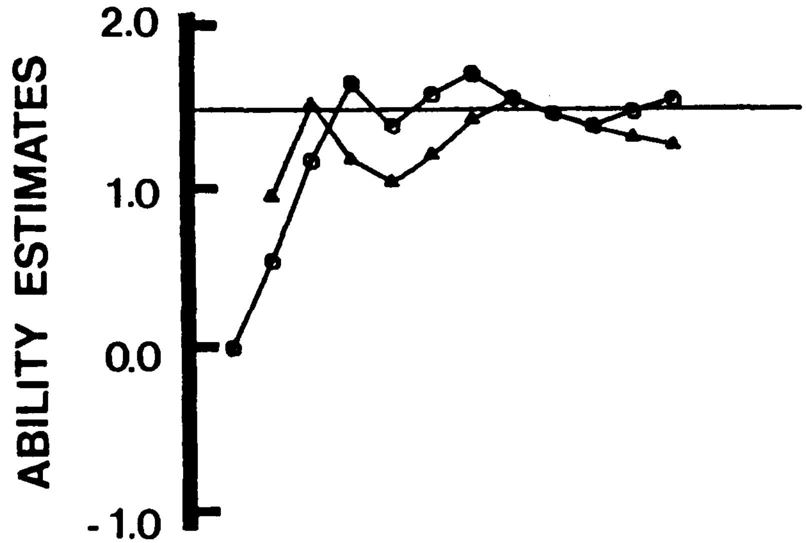

FIGURE 5-1 (Continued): The Prior is n(0,1) , θ = 1.50 .

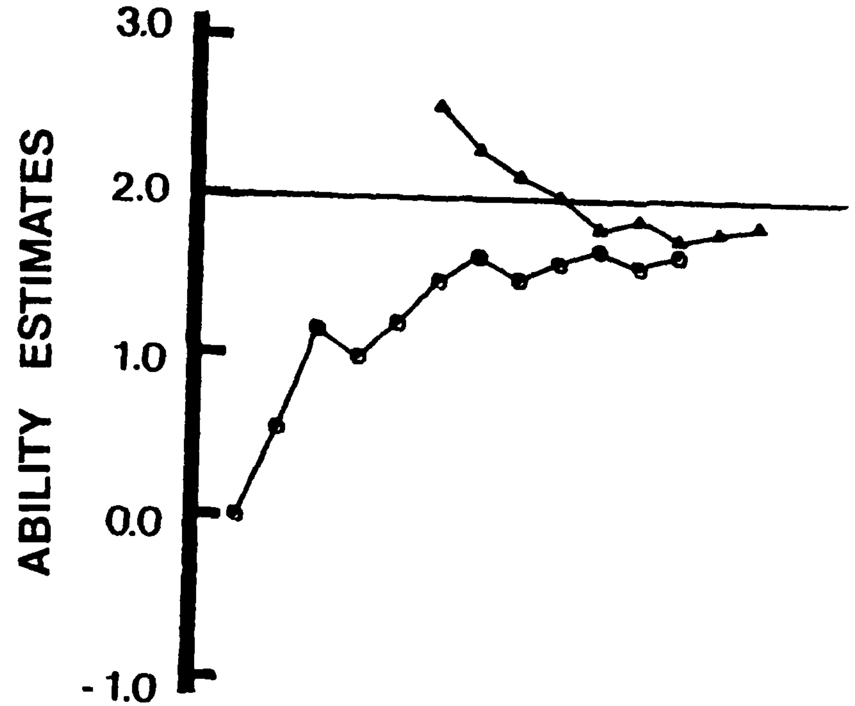

FIGURE 5-1 (Continued): The Prior is n(0,1) , θ = 2.00 .

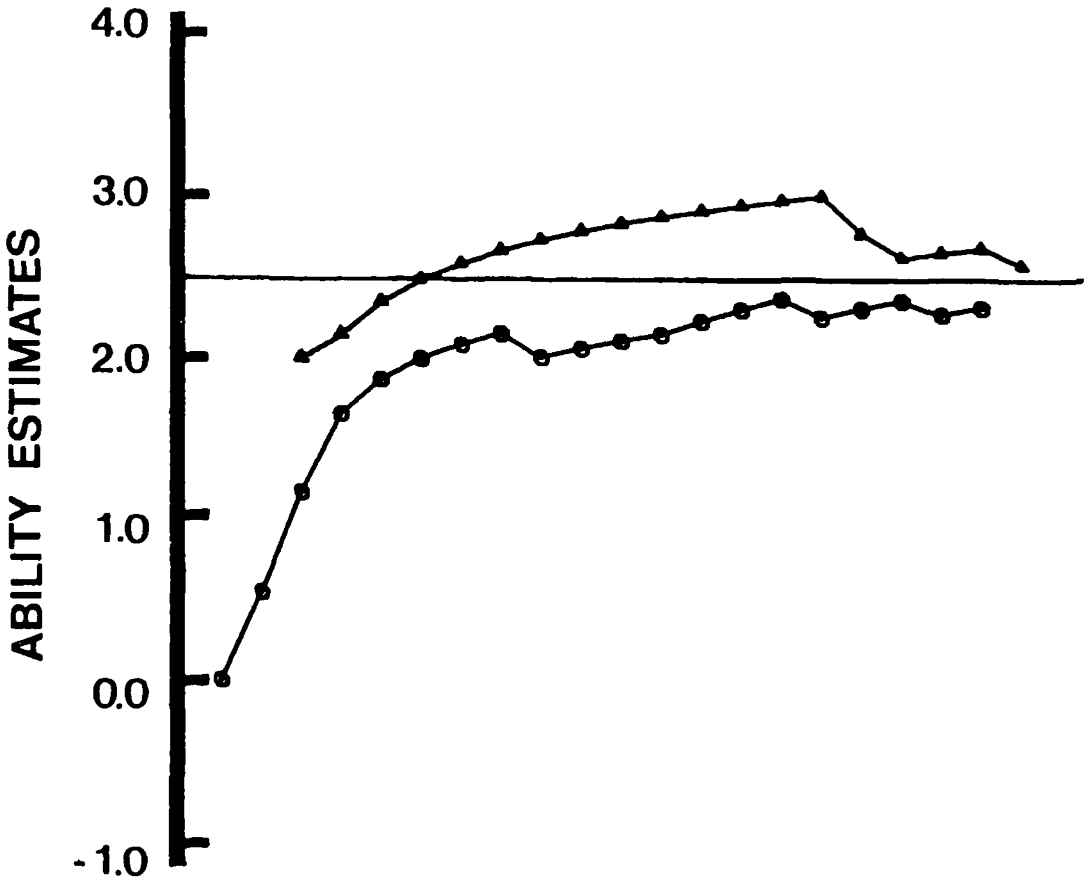

FIGURE 5-1 (Continued): The Prior is n(0,1) , θ = 2.50 .

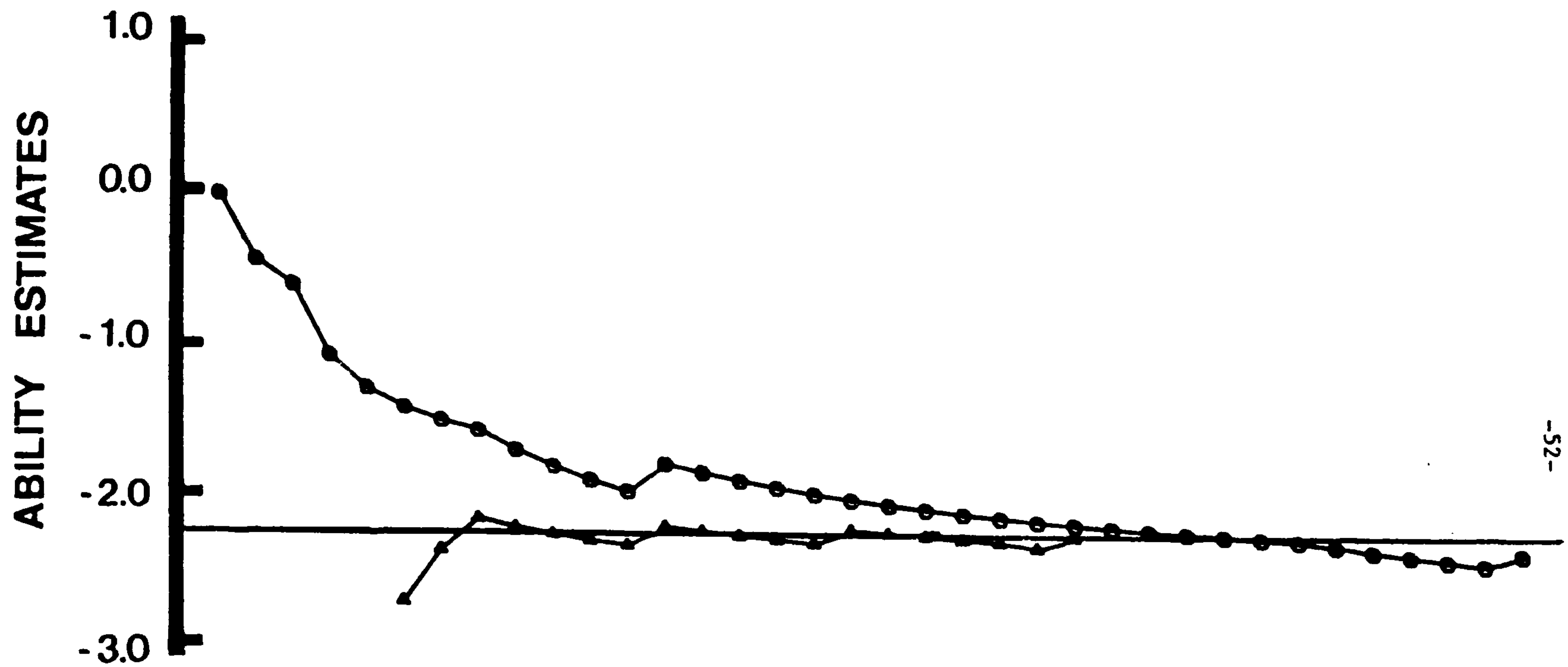

FIGURE 5-2

Successive Maximum Likelihood Estimates (Triangles) and Bayes Modal Estimates (Circles) in the Simulated Tailored Testing with n(0.0,0.8) as the Prior for a Hypothetical Examinee Whose Ability Level is -2.25 .

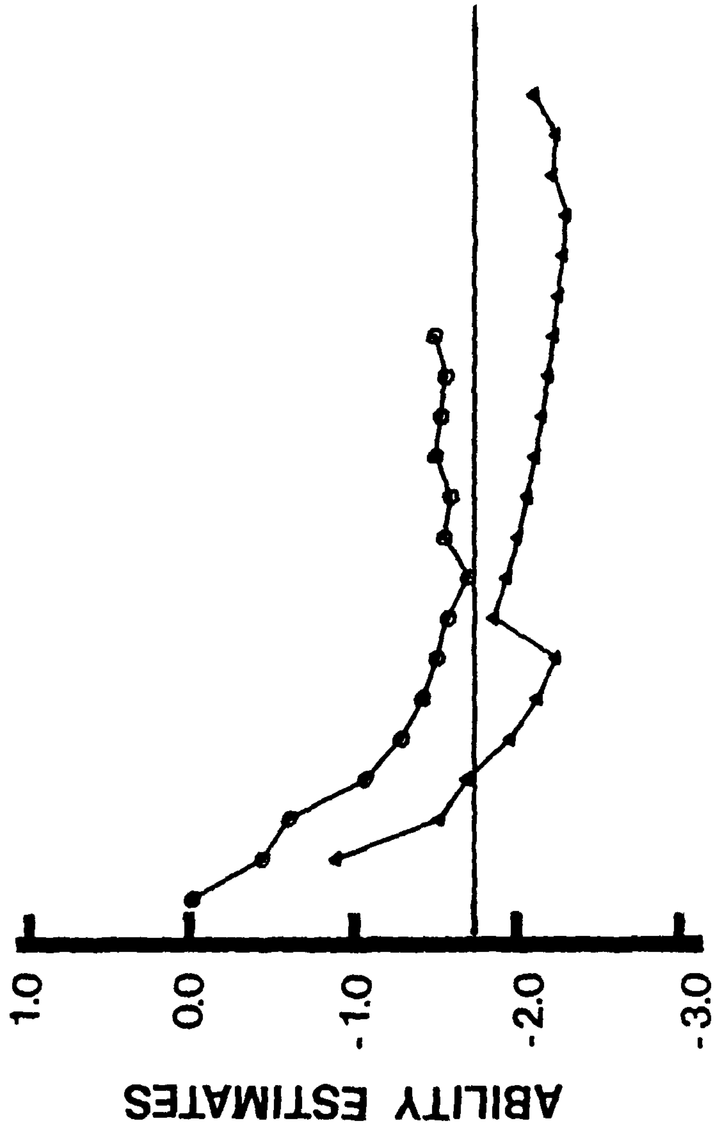

FIGURE 5-2 (Continued): The Prior is n(0.0,0.8) , θ = -1.75 .

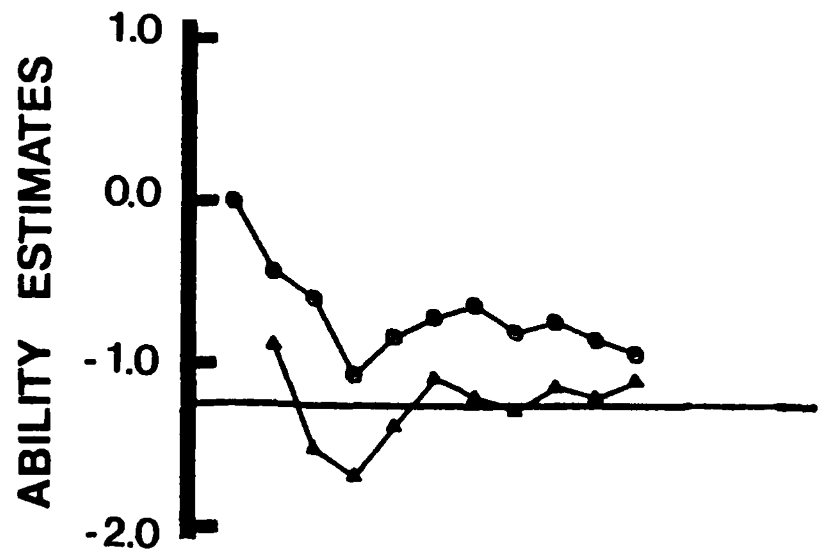

FIGURE 5-2 (Continued): The Prior is n(0.0,0.8) , θ = -1.25 .

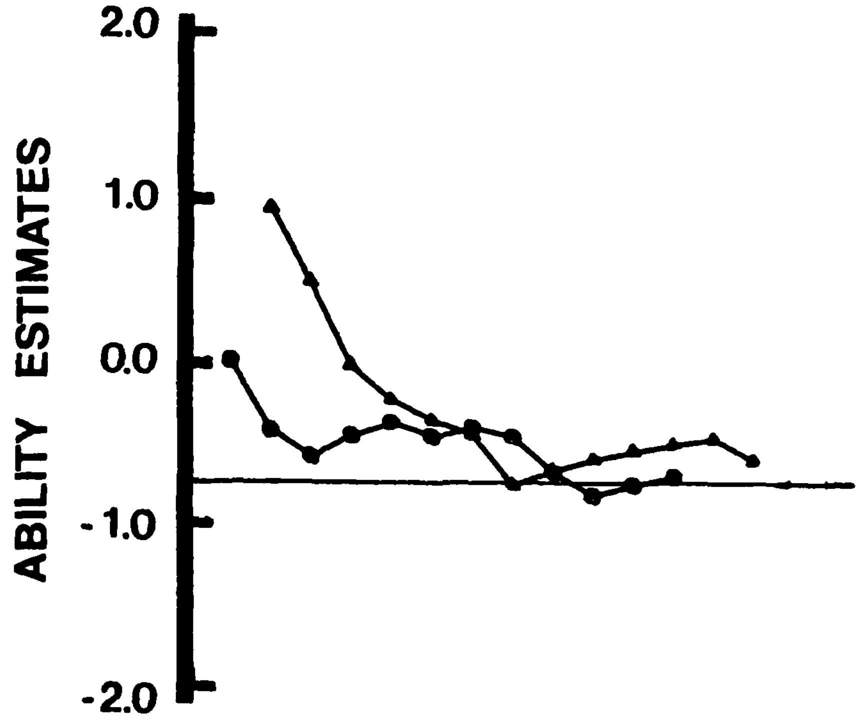

FIGURE 5-2 (Continued): The Prior is n(0.0,0.8) , θ = -0.75 .

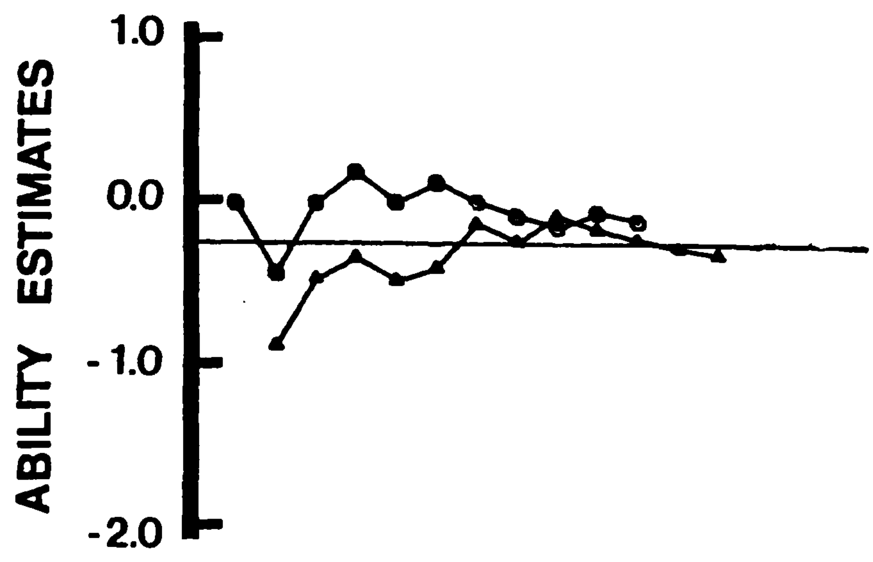

FIGURE 5-2 (Continued): The Prior is n(0.0,0.8) , $\theta = -0.25$.

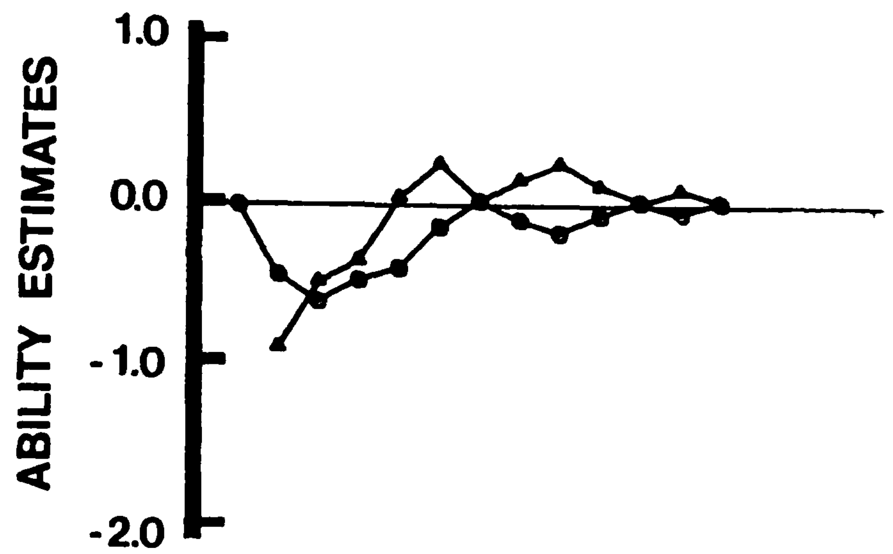

FIGURE 5-2 (Continued): The Prior is n(0.0,0.8) , θ = 0.00 .

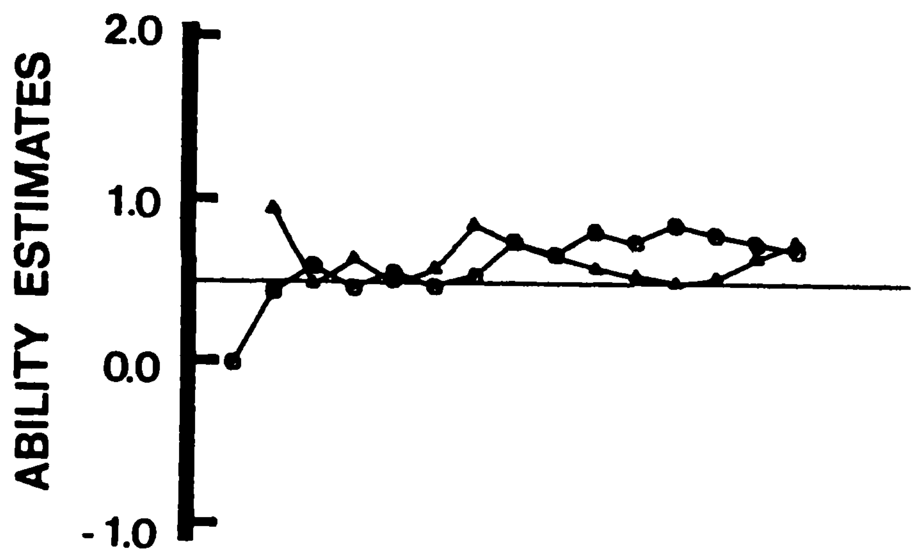

FIGURE 5-2 (Continued): The Prior is n(0.0,0.8) , θ = 0.50 .

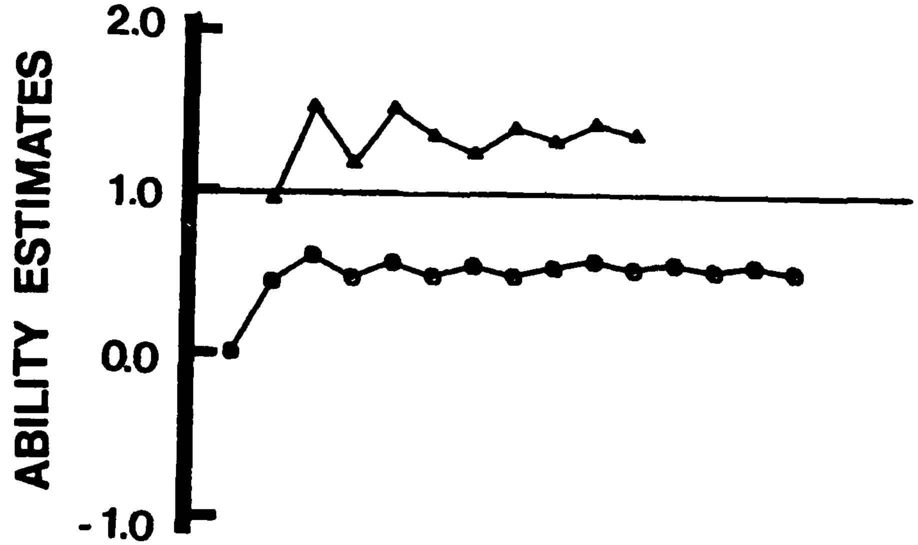

FIGURE 5-2 (Continued): The Prior is n(0.0,0.8) , θ = 1.00 .

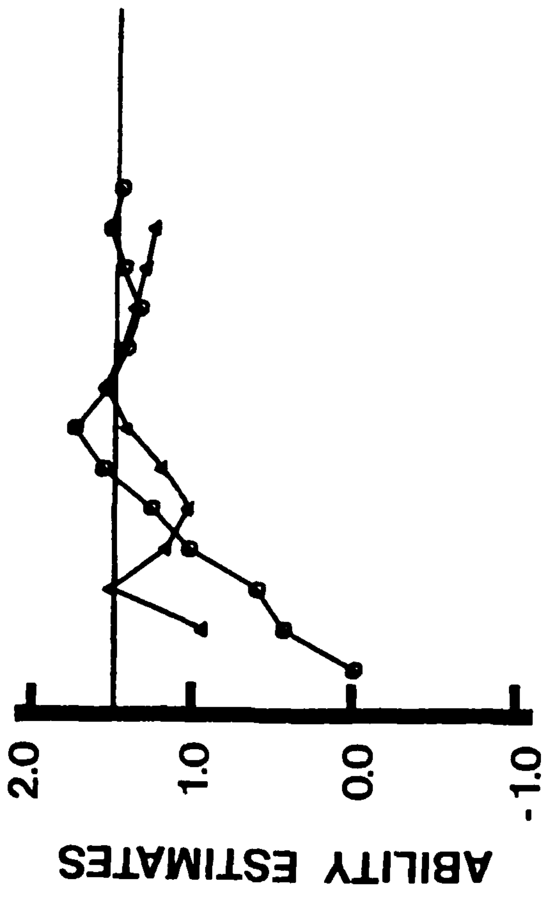

FIGURE 5-2 (Continued): The Prior is n(0.0,0.8) , θ = 1.50 .

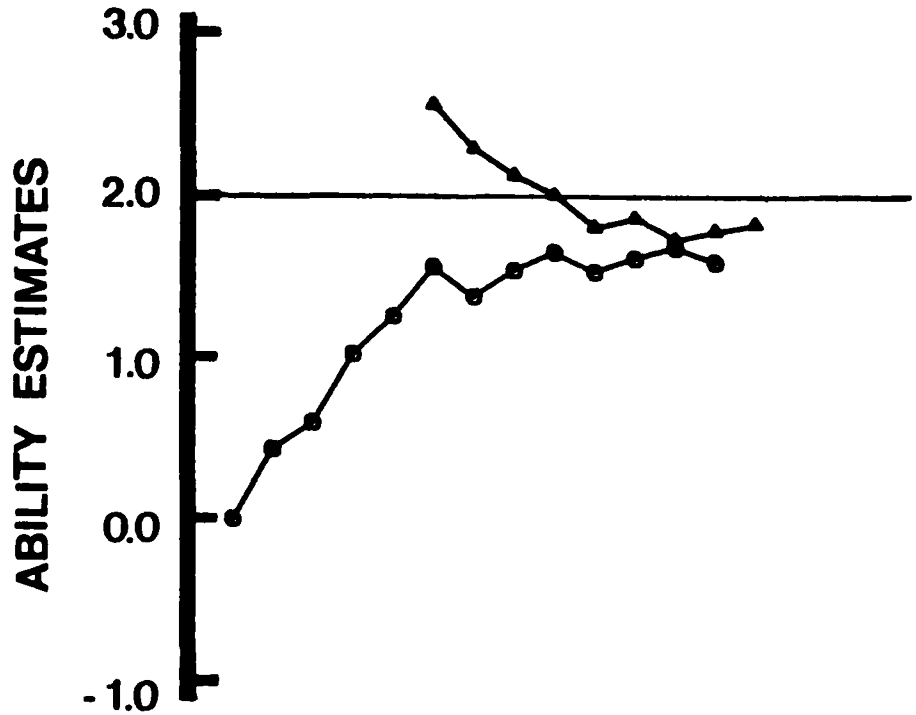

FIGURE 5-2 (Continued): The Prior is n(0.0,0.8) , θ = 2.00 .

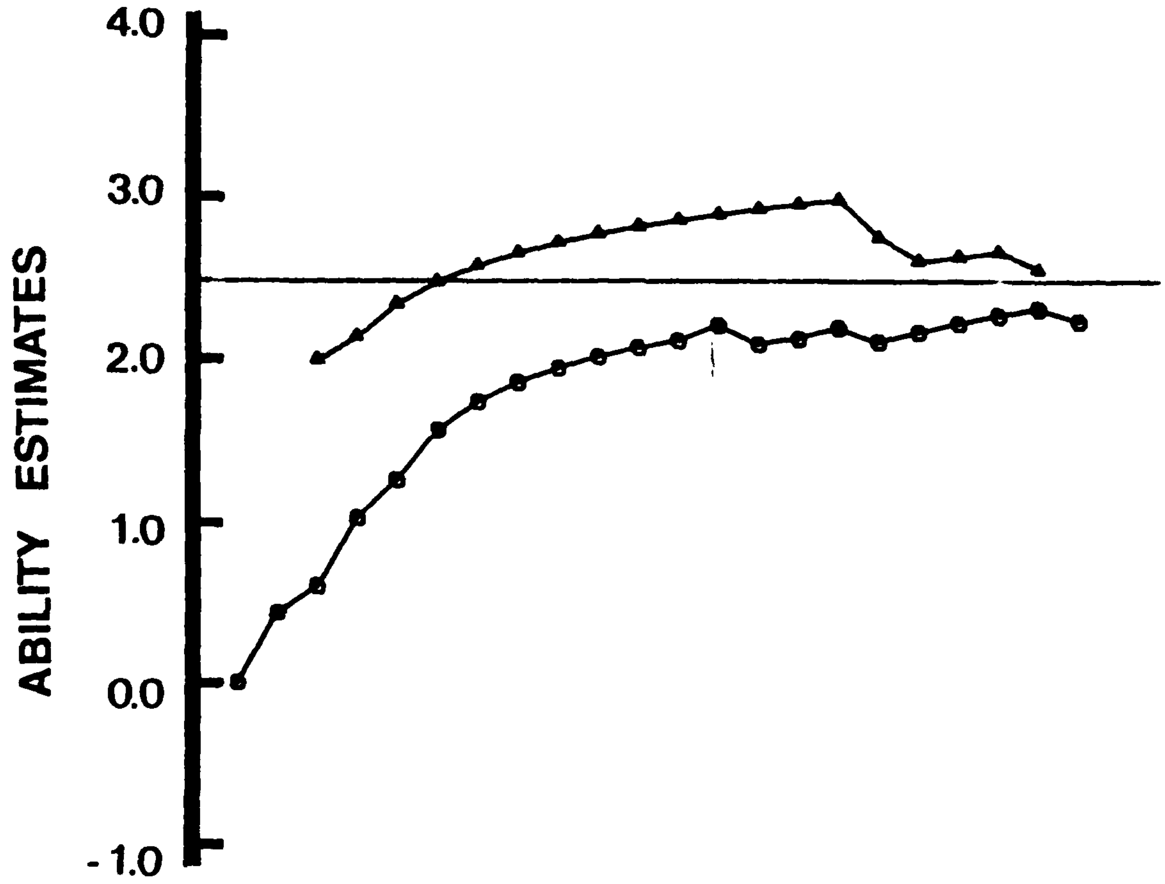

FIGURE 5-2 (Continued): The Prior is n(0.0,0.8) , θ = 2.50 .

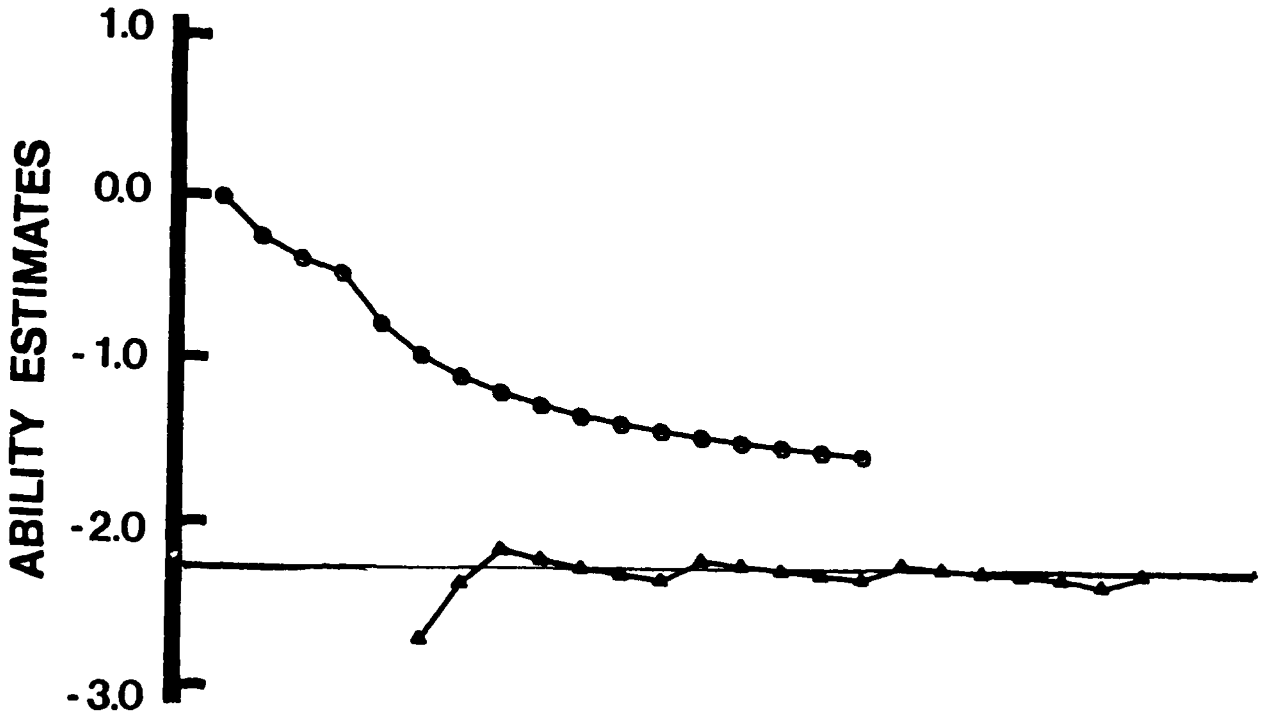

FIGURE 5-3

Successive Maximum Likelihood Estimates (Triangles) and Bayes Modal Estimates (Circles) in the Simulated Tailored Testing with n(0.0,0.5) as the Prior for a Hypothetical Examinee Whose Ability Level is -2.25 .

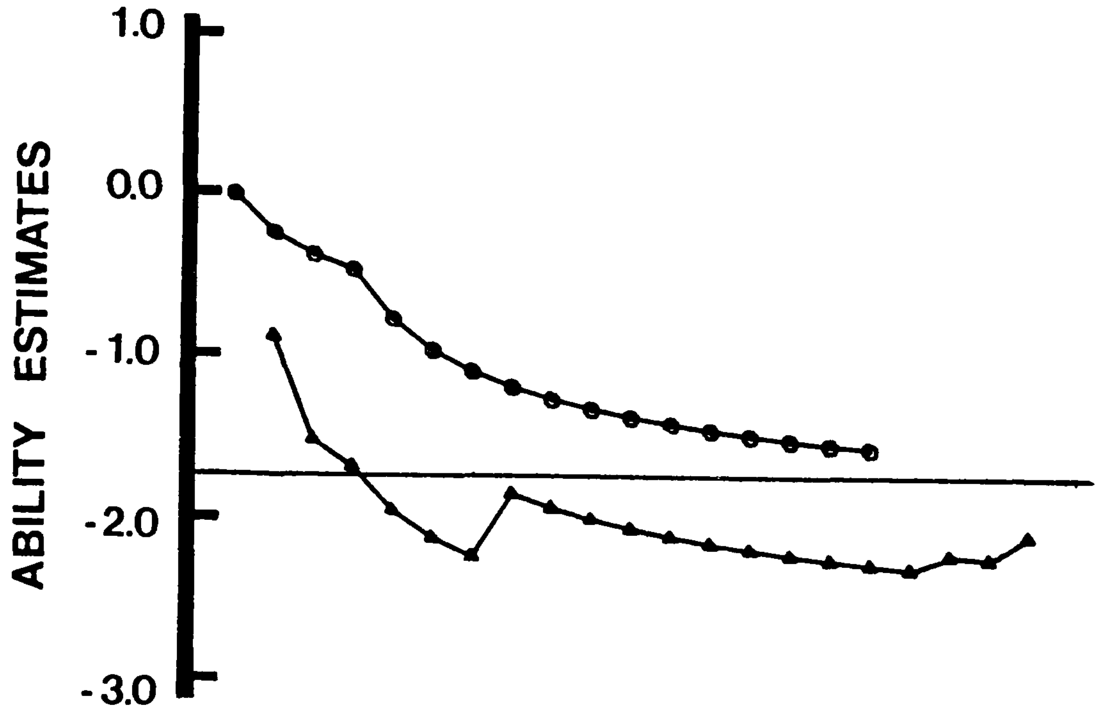

FIGURE 5-3 (Continued): The Prior is n(0.0,0.5) , θ = -1.75 .

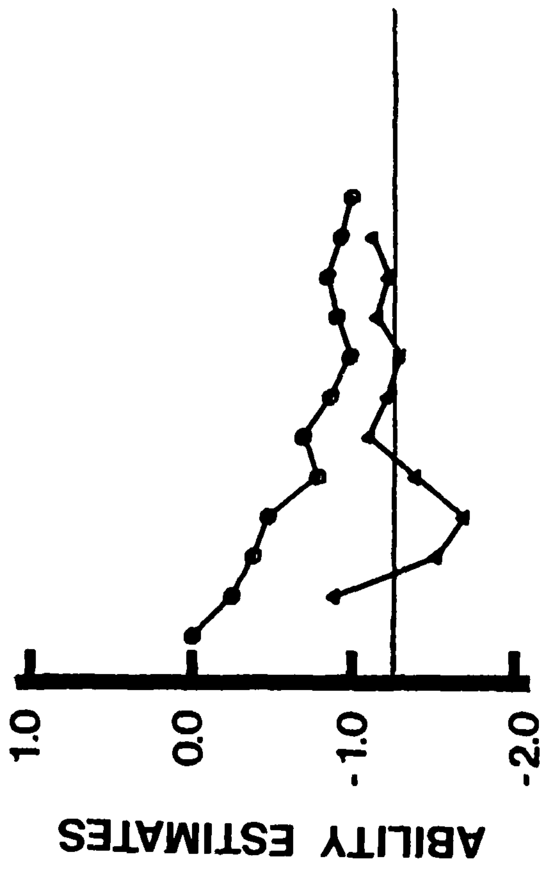

FIGURE 5-3 (Continued): The Prior is n(0.0,0.5) , θ = -1.25 .

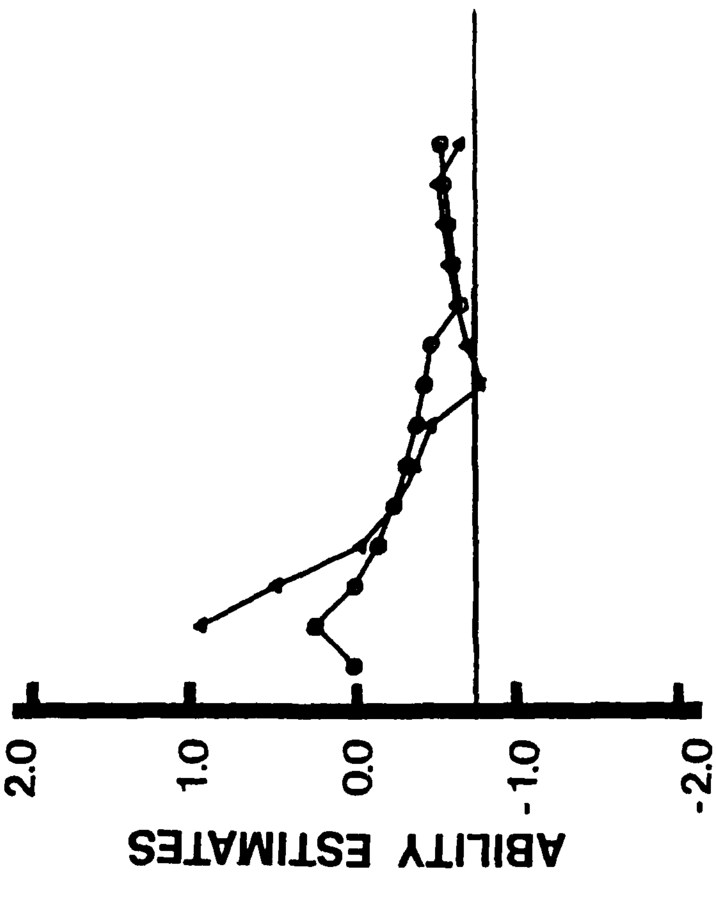

FIGURE 5-3 (Continued): The Prior is n(0.0,0.5) , θ = -0.75 .

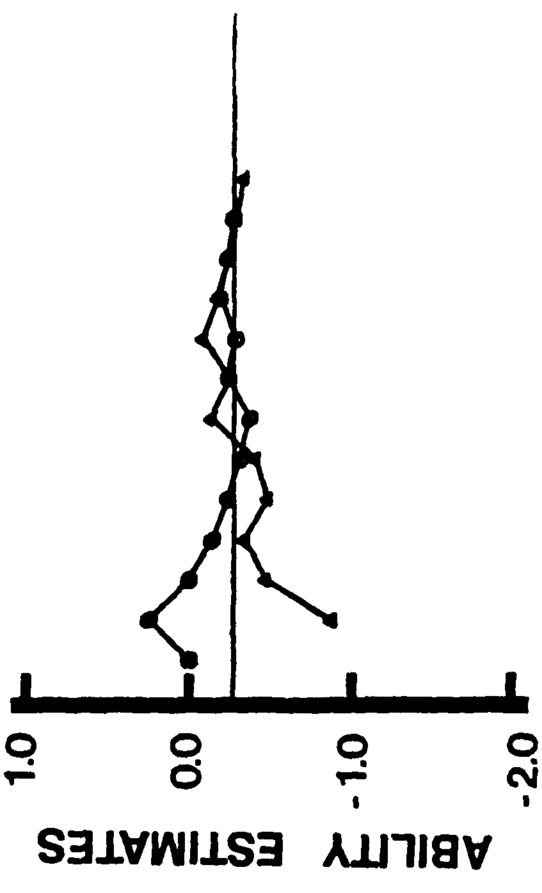

FIGURE 5-3 (Continued): The Prior is n(0.0,0.5) , θ = -0.25 .

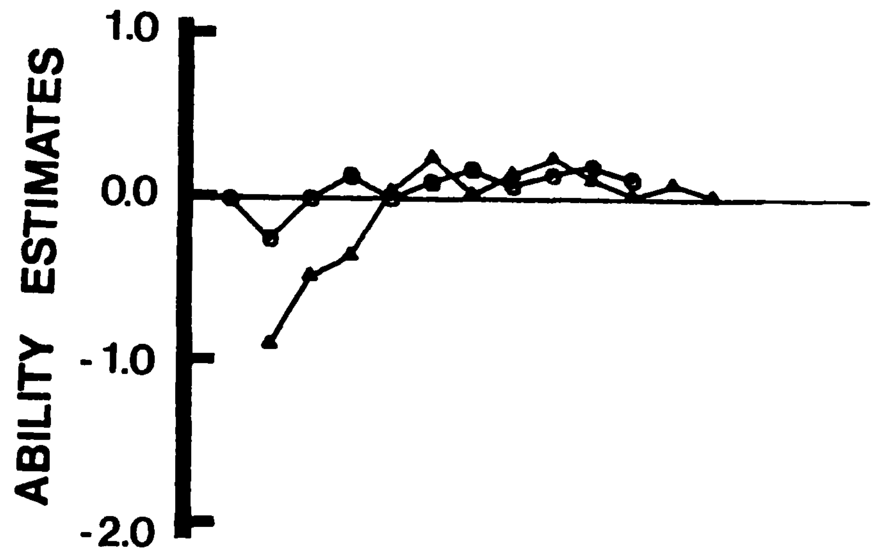

FIGURE 5-3 (Continued): The Prior is n(0.0,0.5) , θ = 0.00 .

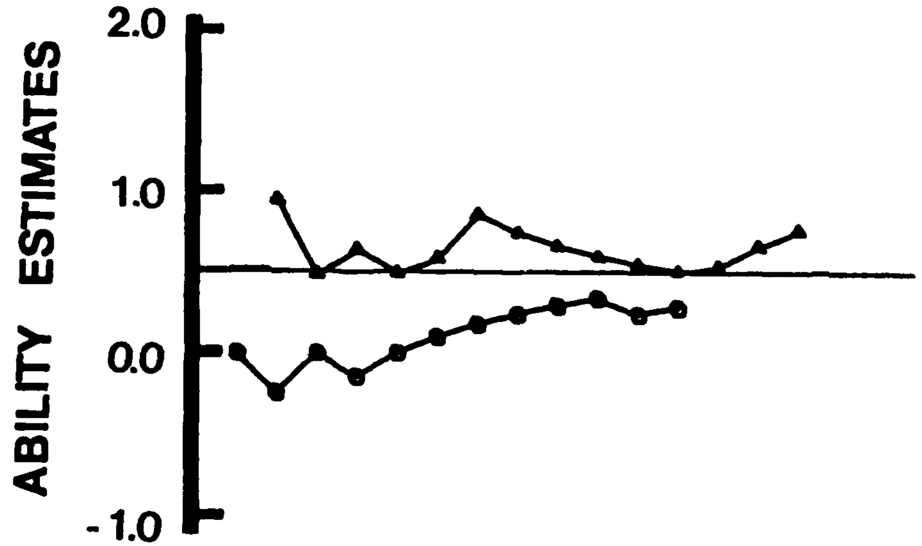

FIGURE 5-3 (Continued): The Prior is n(0.0,0.5) , θ = 0.50 .

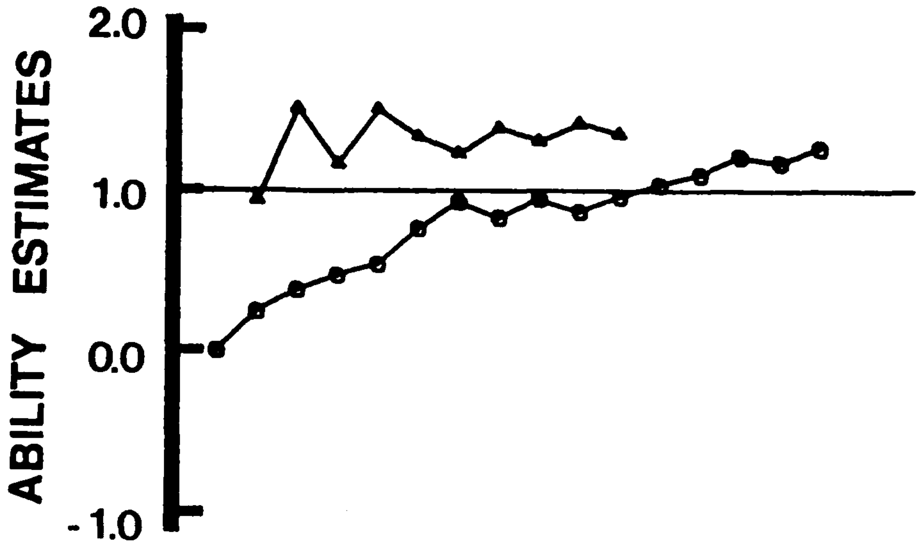

FIGURE 5-3 (Continued): The Prior is n(0.0,0.5) , θ = 1.00 .

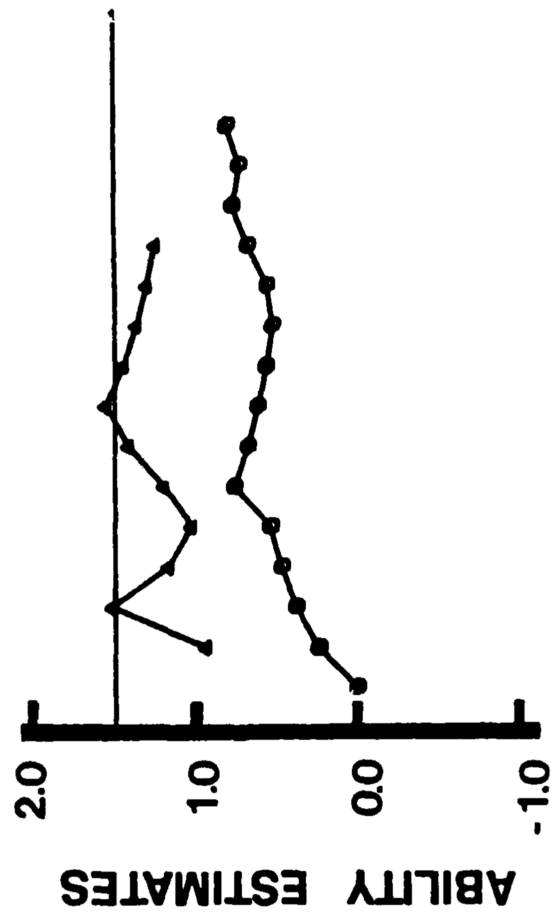

FIGURE 5-3 (Continued): The Prior is n(0.0,0.5) , θ = 1.50 .

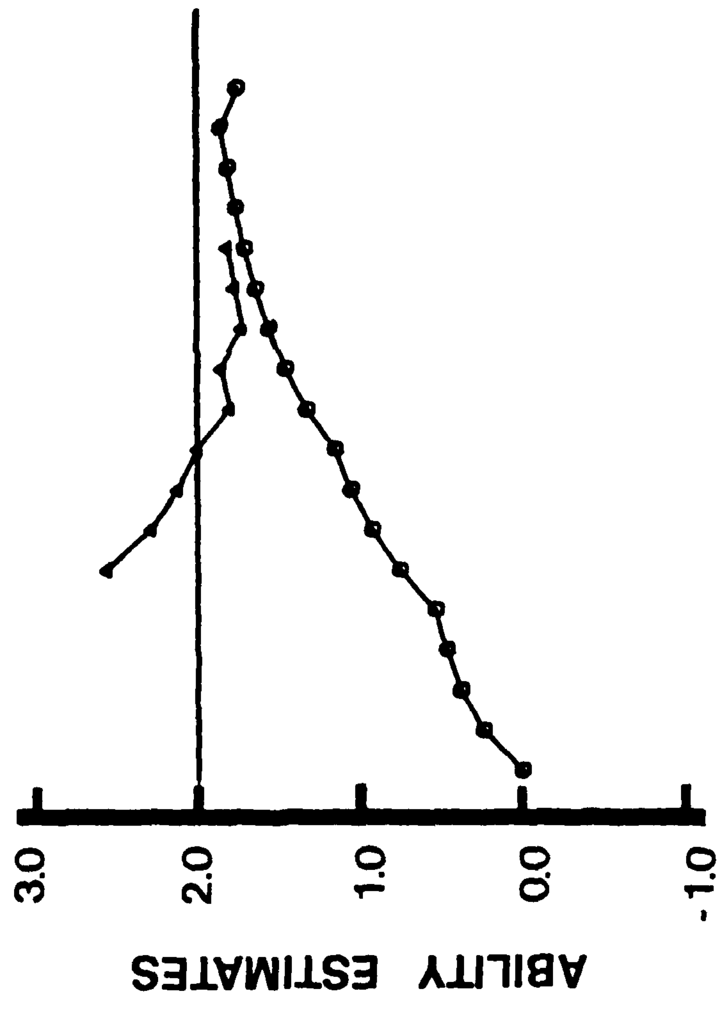

FIGURE 5-3 (Continued): The Prior is n(0.0,0.5) , θ = 2.00 .

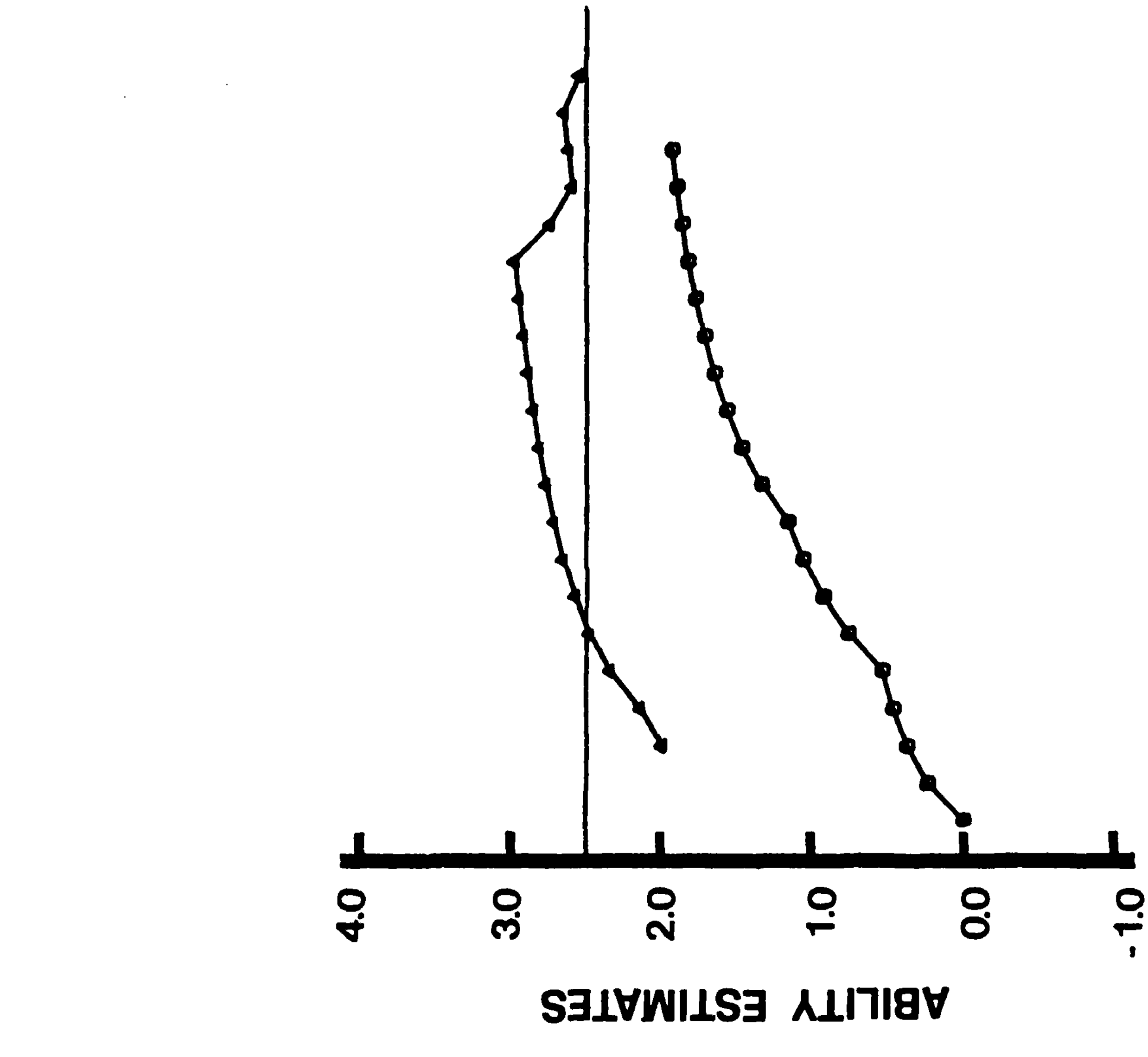

FIGURE 5-3 (Continued): The Prior is n(0.0,0.5) , θ = 2.50 .

TABLE 5-2

Maximum Likelihood Estimate (MLE) and Their Bayes Modal Estimates (BME) of Each of the Eleven Hypothetical Examinees in the Simulated Tailored Testing, with the Numbers of Items Presented in Parentheses, Respectively. The Amount of Test Information Used as the Criterion for the Termination of Presenting New Items Is 20.0 . The Priors Used for the Bayes Modal Estimation Are: (1) n(0.0,1.0) , (2) n(0.0,0.8) , and (3) n(0.0,0.5).

Subject	θ	MLE	BME (1)	BME (2)	BME (3)
1	-2.25	-2.2876 (24)	-2.3161 (28)	-2.4097 (35)	-1.5683 (16)
2	-1.75	-2.1021 (21)	-1.5660 (14)	-1.4873 (14)	-1.5683 (16)
3	-1.25	-1.1120 (11)	-0.7495 (11)	-0.9449 (10)	-0.9859 (11)
4	-0.75	-0.6369 (14)	-0.2581 (13)	-0.7296 (11)	-0.5104 (13)
5	-0.25	-0.3469 (13)	-0.4471 (13)	-0.1306 (10)	-0.2802 (11)
6	0.00	0.0115 (13)	-0.1156 (13)	0.0096 (12)	0.1175 (10)
7	0.50	0.7584 (15)	0.9204 (13)	0.7097 (14)	0.2802 (11)
8	1.00	1.3531 (11)	0.5583 (15)	0.5200 (14)	1.2660 (15)
9	1.50	1.2438 (12)	1.5272 (11)	1.4501 (12)	0.8041 (14)
10	2.00	1.8208 (14)	1.6272 (11)	1.5856 (12)	1.7564 (17)
11	2.50	2.5415 (21)	2.2822 (19)	2.2251 (21)	1.9280 (18)

took 28 and 35 test items to make up for the effect of the priors, n(0.0,1.0) and n(0.0,0.8) , respectively, until the estimates, -2.3161 and -2.4097 , which are slightly less than but close to the true ability level, -2.25 , were obtained. These two numbers of items, 28 and 35 , are far too large compared with the average number of test items used in these forty-four sequences which turned out to be 14.70 .

It is interesting to note that, when the prior is n(0.0,0.5) , it took only 16 items in the sequence, compared with 28 and 35 in the other two Bayesian estimations, before the testing is completed for the examinee whose ability level is -2.25 . The eventual Bayes modal estimate is -1.5683 , however, which is far away from the true ability level -2.25 ; the fact which shows a far stronger effect of the prior than the other two cases. This indicates that the effect of the prior was so strong that the testing could not "correct" the bias caused by the prior throughout the whole process of tailored testing.

A similar tendency of "trade-off" between the accuracy of ability estimation and the number of test items exists in the relationships among the four ability estimates of the same examinee when we lower the criterion for terminating the presentation of a new item. If we use $I(\theta) = 15.0$ as the criterion they are -2.2352 (18) , -2.2796 (21) , -2.0694 (18) and -1.4207 (11) , arranged in the order of the maximum likelihood estimate and Bayes modal estimates with $n(0.0,1.0)$, $n(0.0,0.8)$ and $n(0.0,0.5)$ as the prior, respectively, with the corresponding number of items used in each tailored testing in parenthesis; if we use $I(\theta) = 10.0$ they are -2.2099 (13) , -1.9812 (12) , -1.7915 (12) and -1.1931 (7) . The situation will not be improved even if we switch the criterion to the stability of successive estimates in tailored testing. If, for example, we terminate the presentation of a new item right after three successive estimates stayed within the range of ± 0.075 of the separate, preceding estimates, the resulting four estimates and the numbers of test items for the examinee whose ability level is -2.25 are -2.2994 (11) , -1.9812 (12) , -1.9062 (14) and -1.3781 (10) , respectively. For the examinee of the other deviated ability level, 2.50 , using the same convergence criterion, the results are 2.8183 (11) , 2.1251 (11) , 2.1130 (11) and 1.8152 (15) , compared with those obtained with the criterion of $I(\theta) = 20.0$, which are 2.5415 (21) , 2.2822 (19) , 2.2251 (21) and 1.9280 (18) .

Since only 3 out of 1,000 people are outside of the range of three standard deviations plus or minus the mean, if the population ability distribution is normal, chances are very slim that an examinee whose ability level is -2.25 or 2.50 is assigned to the prior, $n(0.0,0.5)$, or even n(0.0,0.8) . In practice, however, such a situation is more likely to happen, since the function assumed for the prior is more or less arbitrary, and, furthermore, the assignment of individuals to a specific prior itself is more or less arbitrary, using their sex, ethnic background, and so forth. We must say, therefore, Bayesian estimation applied for personnel selection, for example, could cause a serious problem of unfair discrimination, even if the content of the test itself is perfectly valid.

The sequential results of the four ability estimations in the simulated tailored testing for the eleven hypothetical examinees are presented in Appendix as Tables A-1 through A-4.

VI Some Criticisms on the Common Belief in Bayesian Estimation

Many researchers who use Bayesian estimates in favor of the maximum likelihood estimate refer to the two aspects which they think are the advantages of the Bayesian estimation over the maximum likelihood estimation. These two aspects are:

(1) While Bayesian estimation provides us with seemingly reasonable finite values as the estimates for all the response patterns, the maximum likelihood estimation gives us positive and negative infinities for the two extreme response patterns, $(0,0,\ldots,0)'$ and $(m_1,m_2,\ldots,m_n)'$, respectively.

(2) The frequency distribution of the resultant set of maximum likelihood estimates is more scattered than the true ability distribution, or the prior, whereas that of Bayesian estimates is not.

It should be recalled that with each of the two hypothetical tests, i.e., Tests A and B, which were introduced in Chapter 4, every single hypothetical examinee of the two groups obtained a finite maximum likelihood estimate. This results from the fact that the amount of test information of each test is as large as 21.6 for the interval of θ, $(-3.0,3.0)$, the range within which all the examinees' ability is located, and for this reason none of the examinees obtained either of the two extreme response patterns. In fact, it can be proved easily that the conditional probability with which the examinee of a given ability level obtains the response pattern, $(0,0,\dots,0)'$, or $(m_1,m_2,\dots,m_n)'$, approaches zero as the amount of test information increases at that level of θ. It is highly unlikely, therefore, that an examinee whose ability level is located in the interval of θ, $(-3.0,3.0)$, will obtain one of these two extreme response patterns, if the test is highly informative throughout the interval, as is true with both Tests A and B.

If the test is short, like LIS-U, however, it is more likely that examinees obtain one of the two extreme response patterns. This test, LIS-U, was originally developed for the purpose of classifying a group of examinees, whose ability distributes, approximately, normally with zero and unity as the two parameters, into small subgroups of equal sizes (cf. Indow and Samejima, 1962, 1966). The test information function assumes, therefore, high values around $\theta = 0.0$, and lower values as θ departs from zero, as we can see in Figure 3-1, since for the above purpose it is not important to estimate the ability of very bright or very dull examinees accurately. Note that for the classification purpose, negative and positive infinities do not interfere with our process, although as the values of estimates they are far from being acceptable.

It is frequently observed, however, that a researcher uses a test whose test information function is bell-shaped, like the one for LIS-U , for other purposes without giving much thought to the inaccuracy of estimation at deviated values of θ . This inaccuracy is rooted in the test itself, but he tends to blame the method of estimation, complaining that they obtained positive and negative infinities for some examinees as their maximum likelihood estimates of ability θ . When they come across this result, most researchers turn to Bayesian estimation, as if it gave a solution to the problem. We must note, however, that the seemingly acceptable finite estimates for the two extreme response patterns are basically resultant from the prior only, and the test itself is simply powerless in the entire process of estimation; the fact which explains the large differences among the regressions of the Bayes modal estimates on ability θ with different priors, as was observed in Chapter 3. When researchers use a Bayesian estimate in such a situation, therefore, they are *simply covering up the deficiency of the test they chose, pretending as if the test had enough power to estimate the ability accurately*, while the truth is that the amount of test information provided by the test at those levels of ability is so small that no real testing was performed on these levels of ability. It is the researcher who must take the blame for the failure in choosing a right test, not the maximum likelihood estimation.

As for the second aspect, we must be aware of the fact that, as long as there exists some amount of error of estimation, the frequency distribution of the resultant set of estimates _should_ _be_ _expected_ _to_ _have_ _a_ _larger_ _variance_ than the true ability distribution. It has been pointed out (Samejima, 1977c) that for any unbiased estimator, λ , of ability θ , we can write for the variance of λ

$$(6.1) \qquad \mathrm{Var.}(\lambda) = \mathrm{Var.}(\theta) + E[\mathrm{Var.}(\lambda|\theta)] > \mathrm{Var.}(\theta) \quad .$$

Since the maximum likelihood estimate is asymptotically unbiased, (6.1) approximates the relationship between θ and the maximum likelihood estimate, $\hat{\theta}$, when the amount of test information is large enough for the entire range of θ within which the examinees' ability is located. If the conditional expectation of λ , given θ , is constant, i.e., $\mathrm{Var.}(\lambda|\theta) = \sigma^2$, then we can rewrite (6.1) in the form

$$(6.2) \qquad \mathrm{Var.}(\hat{\theta}) = \mathrm{Var.}(\theta) + \sigma^2 \ .$$

When the test information function assumes a large, constant value for the entire interval of θ within which the examinees' true ability is located, i.e., $I(\theta) = C >> 0$ for this interval of θ , the maximum likelihood estimate conditionally distributes approximately normally, given θ , with θ itself and $C^{-1/2}$ as the two parameters. Thus we can write in such a situation

$$\text{(6.3)} \qquad \text{Var.}(\theta) \doteq \text{Var.}(\theta) + C^{-1} > \text{Var.}(\theta) \quad .$$

If we use either Test A or Test B as our test, for example, the sample variance of the maximum likelihood estimate is expected to be approximately 0.046 larger than the population variance of ability θ , regardless of the value of Var.(θ) .

It is evident, therefore, if some estimator of ability θ provides us with an expected sample variance which is the same as the population variance of the true ability θ , there must be a certain bias which makes the resultant estimate regress toward the central tendency of the ability distribution, as we have seen in the Bayes modal estimate in Chapters 3, 4 and 5. We must consider, therefore, that the characteristic of the maximum likelihood estimate described in (2) is a logical result of the asymptotic unbiasedness, whereas the characteristic of Bayesian estimates is a problem, which is caused by its biasedness.

VII Estimates for the Two Extreme Response Patterns

As we have seen in the preceding chapter, the probability is very low that some of our examinees obtain one of the two extreme response patterns, $(0,0,\ldots,0)$ and $(m_1,m_2,\ldots,m_n)$, if we choose a right test. When the test items follow one of the models which satisfy the unique maximum condition (Samejima, 1969, 1972), like the normal ogive and logistic models, the amount of test information for the range of θ in which the examinees' ability is located is a useful measure for the appropriateness of the test. If our test is informative enough for the entire range of ability θ of our interest, as are Tests A and B in the examples in Chapter 4, chances are very slim that some of our examinees obtain one of the extreme response patterns and, consequently, negative or positive infinity for their maximum likelihood estimates.

We must note, however, that even with such tests as Tests A and B and with groups of examinees whose ability distributes within the interval of θ for which the tests are informative, it can still happen, though very rarely, that some examinees obtain negative or positive infinity as their maximum likelihood estimates of ability. Our question is, therefore, if there is any way to avoid such a situation, without losing the perspective of objective testing, which we shall not be able to accomplish by turning to Bayesian estimation.

From the purpose of objective testing, it is obvious that we should find a solution for this problem without using any information which the test itself does not provide. In so doing, we shall make a population-free estimation, in which examinees are solely evaluated from their performances in the testing.

Hereafter, we shall denote the two extreme response patterns, $(0,0,\ldots,0)'$ and $(m_1,m_2\ldots,m_n)'$, by V-min and V-max , respectively. We notice that the operating characteristic $P_{V\text{-min}}(\theta)$ strictly decreases in θ , and $P_{V\text{-max}}(\theta)$ strictly increases in θ , as long as our test items follow a model, or models, like the normal ogive and logistic models. Thus we can conceive of a critical point, θ_c , which satisfies

$$(7.1) \qquad \begin{cases} P_{V\text{-min}}(\theta) \doteq 0 & \text{for } \theta > \theta_c \\ P_{V\text{-max}}(\theta) \doteq 0 & \text{for } \theta \leq \theta_c \,. \end{cases}$$

Let Q be the product of the two operating characteristics of the response patterns, V-min and V-max , such that

$$(7.2) \qquad Q = P_{V\text{-min}}(\theta) \; P_{V\text{-max}}(\theta) \,.$$

We define θ_c as the point of θ at which this product is minimal. By virtue of the assumption of local independence (Lord and Novick, 1968), θ_c is the solution for the equation:

$$(7.3)\quad \frac{\partial}{\partial\theta}\log Q = \frac{\partial}{\partial\theta}\log P_{V\text{-min}}(\theta) + \frac{\partial}{\partial\theta}\log P_{V\text{-max}}(\theta)$$

$$= \sum_{g=1}^{n} \frac{\partial}{\partial\theta}\log P_{x_g}(\theta;x_g=0)$$

$$+ \sum_{g=1}^{n} \frac{\partial}{\partial\theta}\log P_{x_g}(\theta;x_g=m_g)$$

$$= \sum_{g=1}^{n} A_{x_g}(\theta;x_g=0) + \sum_{g=1}^{n} A_{x_g}(\theta;x_g=m_g)$$

$$= 0 ,$$

where $A_{x_g}(\theta)$ is the basic function (Samejima, 1969, 1972) of the item response x_g . It is interesting to note that this critical value θ_c is the maximum likelihood estimate of ability θ for the response pattern, $(0,0,\ldots,0,1,1,\ldots,1)$, on the test of $2n$ binary items, the first n items of which have $P_{x_g}(\theta;x_g=1)$ $(g=1,2,\ldots,n)$, and the second n items of which have $P_{x_g}(\theta;x_g=m_g)$ $(g=1,2,\ldots,n)$, as their respective item characteristic functions.

We shall aim at finding finite substitutes for the two maximum likelihood estimates, $\hat{\theta}_{V\text{-min}}$ and $\hat{\theta}_{V\text{-max}}$, which are negative and positive infinities, respectively, in such a way that the substitution should provide us with a regression which is close enough to θ , i.e., the unbiasedness of the estimator, for some range of θ . Let $\theta^*_{V\text{-min}}$ and $\theta^*_{V\text{-max}}$ denote such estimates, and θ^* be the resultant estimator, such that

$$(7.4)\qquad \theta^* \begin{cases} = \theta^*_{V\text{-min}} & \text{for } V\text{-min} \\ = \theta^*_{V\text{-max}} & \text{for } V\text{-max} \\ = \hat{\theta}_V & \text{for all the other response patterns.} \end{cases}$$

We can write for the regression of θ^* on ability θ such that

$$(7.5)\qquad E(\theta^*|\theta) = \sum_{\substack{V\neq V\text{-min}\\ V\neq V\text{-max}}} \hat{\theta}_V\ P_V(\theta) + \theta^*_{V\text{-min}}\ P_{V\text{-min}}(\theta) + \theta^*_{V\text{-max}}\ P_{V\text{-max}}(\theta)$$

$$\begin{cases} \doteq \sum\limits_{\substack{V\neq V\text{-min}\\ V\neq V\text{-max}}} \hat{\theta}_V\ P_V(\theta) + \theta^*_{V\text{-min}}\ P_{V\text{-min}}(\theta) & \text{for } \theta \leq \theta_c \\ \doteq \sum\limits_{\substack{V\neq V\text{-min}\\ V\neq V\text{-max}}} \hat{\theta}_V\ P_V(\theta) + \theta^*_{V\text{-max}}\ P_{V\text{-max}}(\theta) & \text{for } \theta > \theta_c\ . \end{cases}$$

If this estimator, θ^* , provides us with an approximate unbiasedness for a certain range of θ , $(\underline{\theta}, \bar{\theta})$, then we shall be able to write

$$(7.6)\qquad \begin{cases} \sum\limits_{\substack{V\neq V\text{-min}\\ V\neq V\text{-max}}} \hat{\theta}_V\ P_V(\theta) + \theta^*_{V\text{-min}}\ P_{V\text{-min}}(\theta) \doteq \theta & \text{for } \underline{\theta} < \theta \leq \theta_c \\ \sum\limits_{\substack{V\neq V\text{-min}\\ V\neq V\text{-max}}} \hat{\theta}_V\ P_V(\theta) + \theta^*_{V\text{-max}}\ P_{V\text{-max}}(\theta) \doteq \theta & \text{for } \theta_c < \theta < \bar{\theta}\ . \end{cases}$$

In practice, we must search the interval of θ , $(\underline{\theta}, \bar{\theta})$, for which such an estimator, θ^* , is availiable, in relation with a specific test of our interest. From (7.6), we can further write

$$
(7.7)\quad \begin{cases} \sum_{\substack{V \neq V\text{-min} \\ V \neq V\text{-max}}} \hat{\theta}_V \int_{\underline{\theta}}^{\theta_c} P_V(\theta)\, d\theta + \theta^*_{V\text{-min}} \int_{\underline{\theta}}^{\theta_c} P_{V\text{-min}}(\theta)\, d\theta \\ \qquad \doteq \frac{1}{2}(\theta_c^2 - \bar{\theta}^2) \\ \sum_{\substack{V \neq V\text{-min} \\ V \neq V\text{-max}}} \hat{\theta}_V \int_{\theta_c}^{\bar{\theta}} P_V(\theta)\, d\theta + \theta^*_{V\text{-max}} \int_{\theta_c}^{\bar{\theta}} P_{V\text{-max}}(\theta)\, d\theta \\ \qquad \doteq \frac{1}{2}(\bar{\theta}^2 - \theta_c^2) \ . \end{cases}
$$

Thus the two estimates, $\theta^*_{V\text{-min}}$ and $\theta^*_{V\text{-max}}$, can be obtained by

$$
(7.8)\quad \begin{cases} \theta^*_{V\text{-min}} = [\frac{1}{2}(\theta_c^2 - \underline{\theta}^2) - \sum_{\substack{V \neq V\text{-min} \\ V \neq V\text{-max}}} \hat{\theta}_V \int_{\underline{\theta}}^{\theta_c} P_V(\theta)\, d\theta] \\ \qquad [\int_{\underline{\theta}}^{\theta_c} P_{V\text{-min}}(\theta)\, d\theta]^{-1} \\ \theta^*_{V\text{-max}} = [\frac{1}{2}(\bar{\theta}^2 - \theta_c^2) - \sum_{\substack{V \neq V\text{-min} \\ V \neq V\text{-max}}} \hat{\theta}_V \int_{\theta_c}^{\bar{\theta}} P_V(\theta)\, d\theta] \\ \qquad [\int_{\theta_c}^{\bar{\theta}} P_{V\text{-max}}(\theta)\, d\theta]^{-1} \ , \end{cases}
$$

with some appropriate values for $\underline{\theta}$ and $\bar{\theta}$.

For the purpose of illustration, we use LIS-U again, and put our effort upon finding a suitable interval, $(\underline{\theta}, \bar{\theta})$, and the corresponding two estimates, $\theta^*_{V\text{-min}}$ and $\theta^*_{V\text{-max}}$, which substitute for negative and positive infinities, respectively, in the maximum likelihood estimation. We must be aware that this short test, LIS-U, is not designed for estimating a wide range of ability θ with high accuracy, as was explained in the preceding chapter. This implies that we should expect the interval $(\underline{\theta}, \bar{\theta})$ to be a relatively small one, for which the test information function, $I(\theta)$, assumes reasonably high values (cf. Figure 3-1.)

We obtained -0.00880 for the critical value, θ_c , for LIS-U, which is the solution for (7.3). For the endpoints of the interval, $\underline{\theta}$ and $\bar{\theta}$, we used eleven different sets, ± 1.50 , ± 1.75 , ± 2.00 , ± 2.25 , ± 2.50 , ± 3.00 , ± 3.50 , ± 4.00 , ± 4.50 , ± 5.00 and ± 5.50 , for the purpose of experimentation. The resultant set of estimates, $\theta^*_{V\text{-min}}$ and $\theta^*_{V\text{-max}}$, which was obtained by using each of these eleven intervals, is given in Table 7-1. We can see that the value of $\theta^*_{V\text{-min}}$ decreases as the lower endpoint of the interval $\underline{\theta}$, decreases, and that of $\theta^*_{V\text{-max}}$ increases as the upper endpoint, $\bar{\theta}$, increases, as is expected from (7.8). In order to find out if (7.1) is satisfied, the two quantities, S_L and S_U , such that

$$(7.9)\quad \begin{cases} S_L = \sum_{V \neq V\text{-max}} \int_{\underline{\theta}}^{\theta_c} P_V(\theta)\, d\theta \\ S_U = \sum_{V \neq V\text{-min}} \int_{\theta_c}^{\bar{\theta}} P_V(\theta)\, d\theta \end{cases}$$

TABLE 7-1

Eleven Sets of Estimates, $\theta^*_{V\text{-min}}$ and $\theta^*_{V\text{-max}}$, of Ability for the Two Extreme Response Patterns, (0,0,...,0) and (1,1,...,1), Obtained on LIS-U, Using Eleven Different Intervals for $(\underline{\theta},\bar{\theta})$.

$\underline{\theta}$, $\bar{\theta}$	$\theta^*_{V\text{-min}}$	$\theta^*_{V\text{-max}}$
± 1.50	-1.47883	1.52237
± 1.75	-1.64702	1.65605
± 2.00	-1.79255	1.77649
± 2.25	-1.92540	1.89233
± 2.50	-2.05136	2.00754
± 3.00	-2.29490	2.24127
± 3.50	-2.53641	2.48011
± 4.00	-2.77945	2.72254
± 4.50	-3.02430	2.96720
± 5.00	-3.27051	3.21329
± 5.50	-3.51765	3.46032

were computed, and presented in Table 7-2. We can see that these values are very close to the areas, which are obtainable if we include all the response patterns, and are equal to $(\theta_c - \underline{\theta})$ and $(\bar{\theta} - \theta_c)$, respectively; in fact, the discrepancies of S_L and S_U from these values are approximately 0.00037 and 0.00035 , respectively, in each of the eleven cases, the fact which indicates the satisfaction of (7.1).

The regression of θ^* on ability θ , which is given in the first two lines of (7.5), was computed by using each of the eleven sets of $\theta^*_{V\text{-min}}$ and $\theta^*_{V\text{-max}}$, and the first five cases are presented as Figure 7-1. In each of these five graphs, the regression, $E(\theta^*|\theta)$, is drawn by a solid curve, and, for the sake of comparison, the regression of the Bayes modal estimate with $n(0,1)$ as the prior is plotted by dots, together with the solid straight line which indicates the unbiasedness. We can see in these results that all the regressions of θ^* on θ are closer to the unbiasedness than the regression of the Bayes modal estimate, and, in fact, for the first three cases in which the interval, $(\underline{\theta}, \bar{\theta})$, is $(-1.50,1.50)$, $(-1.75,1.75)$ and $(-2.00,2.00)$, respectively, $E(\theta^*|\theta)$ is very close to the straight line within the respective intervals, $(\underline{\theta}, \bar{\theta})$. The departure of $E(\theta^*|\theta)$ from the unbiasedness becomes greater as we change the interval to $(-2.25,2.25)$ and $(-2.50,2.50)$, the result which was anticipated from the test information function

TABLE 7-2

Sum of the Areas, S_L , Under the Curves of $P_V(\theta)$, Excluding $P_{V-max}(\theta)$, for the Interval $(\underline{\theta}, \theta_c)$, and the Sum of the Areas, S_U , Under $P_V(\theta)$, Excluding $P_{V-min}(\theta)$, for the Interval, $(\theta_c, \overline{\theta})$, Together With Their Sum.

$\underline{\theta}$, $\overline{\theta}$	S_L	S_U	Total
± 1.50	1.49083	1.50845	2.99928
± 1.75	1.74083	1.75845	3.49928
± 2.00	1.99083	2.00845	3.99928
± 2.25	2.24083	2.25845	4.49928
± 2.50	2.49083	2.50845	4.99928
± 3.00	2.99083	3.00845	5.99928
± 3.50	3.49083	3.50845	6.99928
± 4.00	3.99083	4.00845	7.99928
± 4.50	4.49083	4.50845	8.99928
± 5.00	4.99083	5.00845	9.99928
± 5.50	5.49083	5.50845	10.99928

AD-A087 829 TENNESSEE UNIV KNOXVILLE DEPT OF PSYCHOLOGY F/G 12/1

IS BAYESIAN ESTIMATION PROPER FOR ESTIMATING THE INDIVIDUAL'S A--ETC(U)

JUL 80 F SAMEJIMA N00014-77-C-0360

UNCLASSIFIED NL

2 OF 2

AD A 087829

END

DATE FILMED

9-80

DTIC

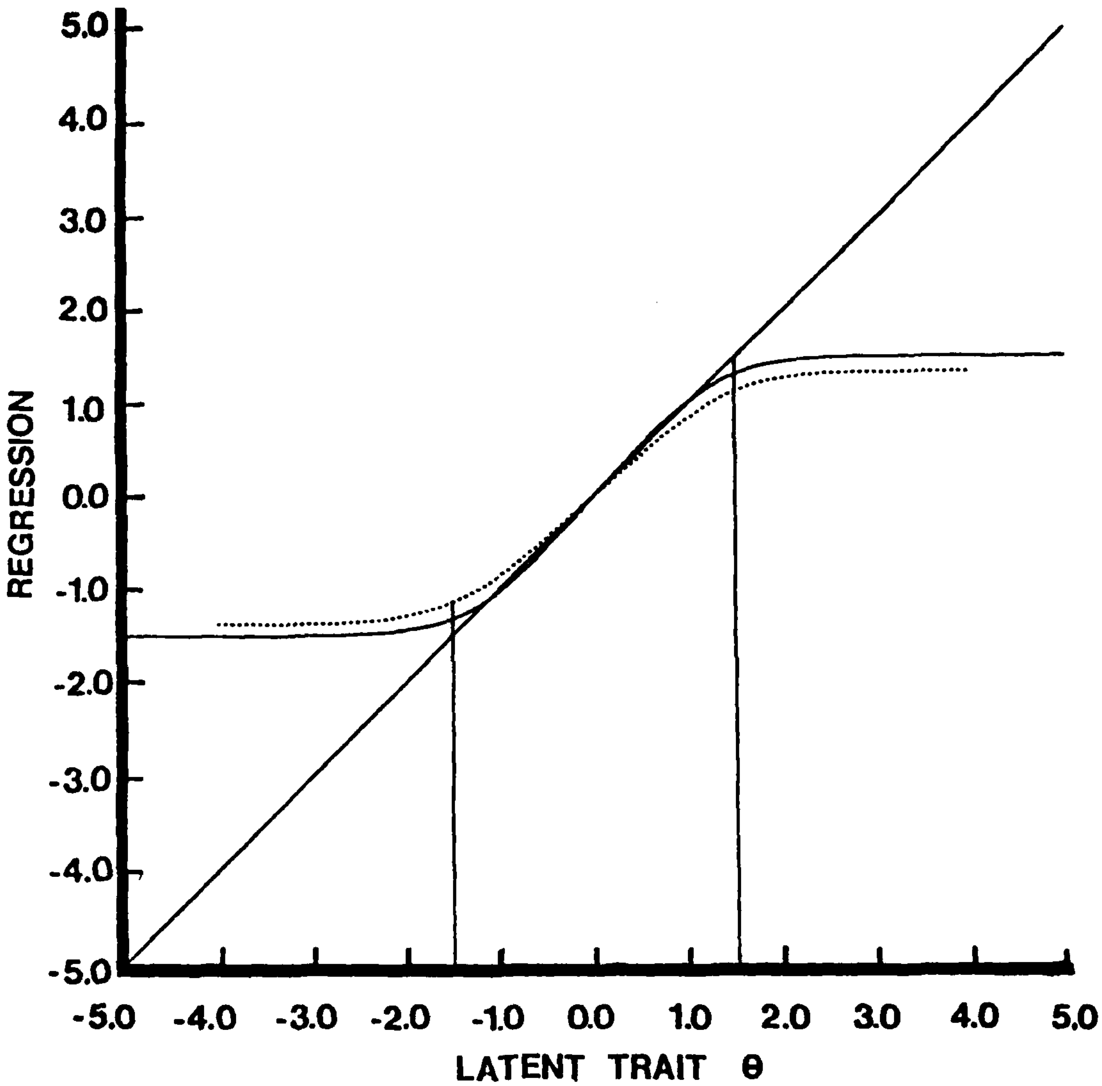

FIGURE 7-1

Regression of the Maximum Likelihood Estimate with Those for the Two Extreme Response Patterns, (0,0,...,0) and (1,1,...,1), Replaced by θ^*_{V-min} and θ^*_{V-max}, Respectively, on Ability θ (Solid Curve), Together with the Regression of the Bayes Modal Estimate with n(0,1) as the Prior (Dotted Curve). These Two Estimates Were Obtained by Using $\underline{\theta} = -1.50$ and $\bar{\theta} = 1.50$.

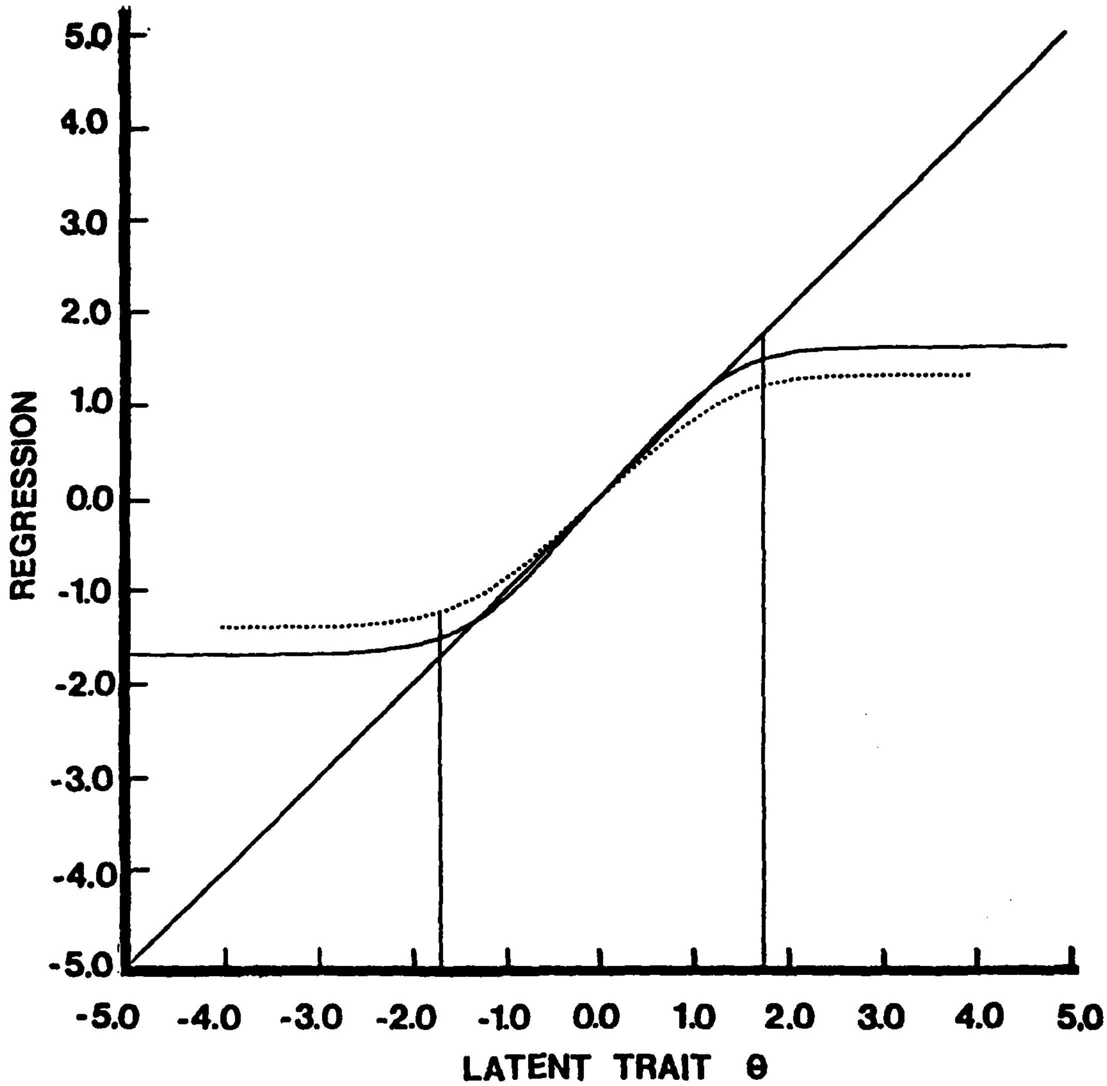

FIGURE 7-1 (Continued)

$\theta^*_{V\text{-}min}$ and $\theta^*_{V\text{-}max}$ Were Obtained by Using $\underline{\theta} = -1.75$ and $\bar{\theta} = 1.75$.

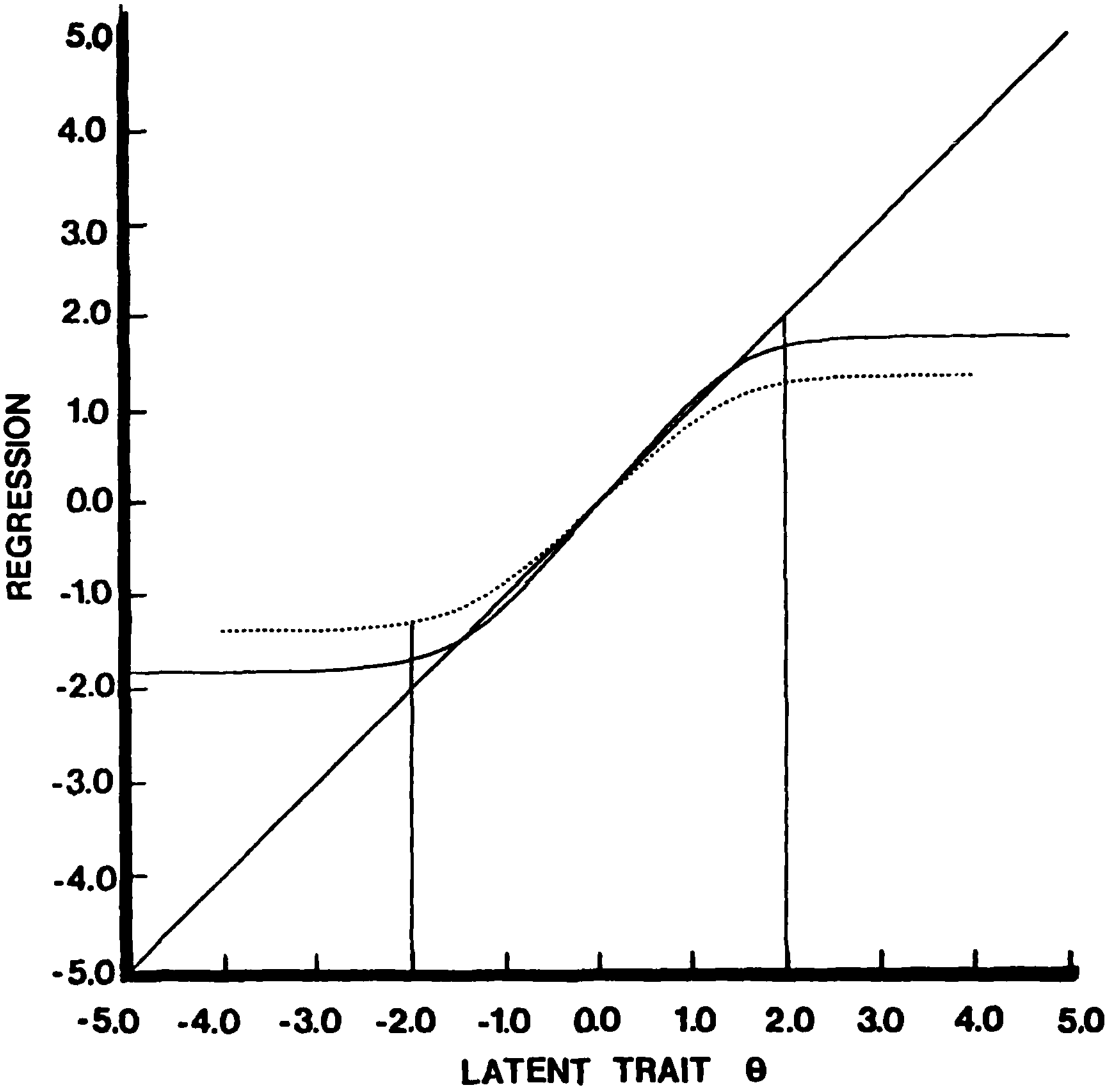

FIGURE 7-1 (Continued)

$\theta^*_{V\text{-min}}$ and $\theta^*_{V\text{-max}}$ Were Obtained by Using $\underline{\theta} = -2.00$ and $\bar{\theta} = 2.00$.

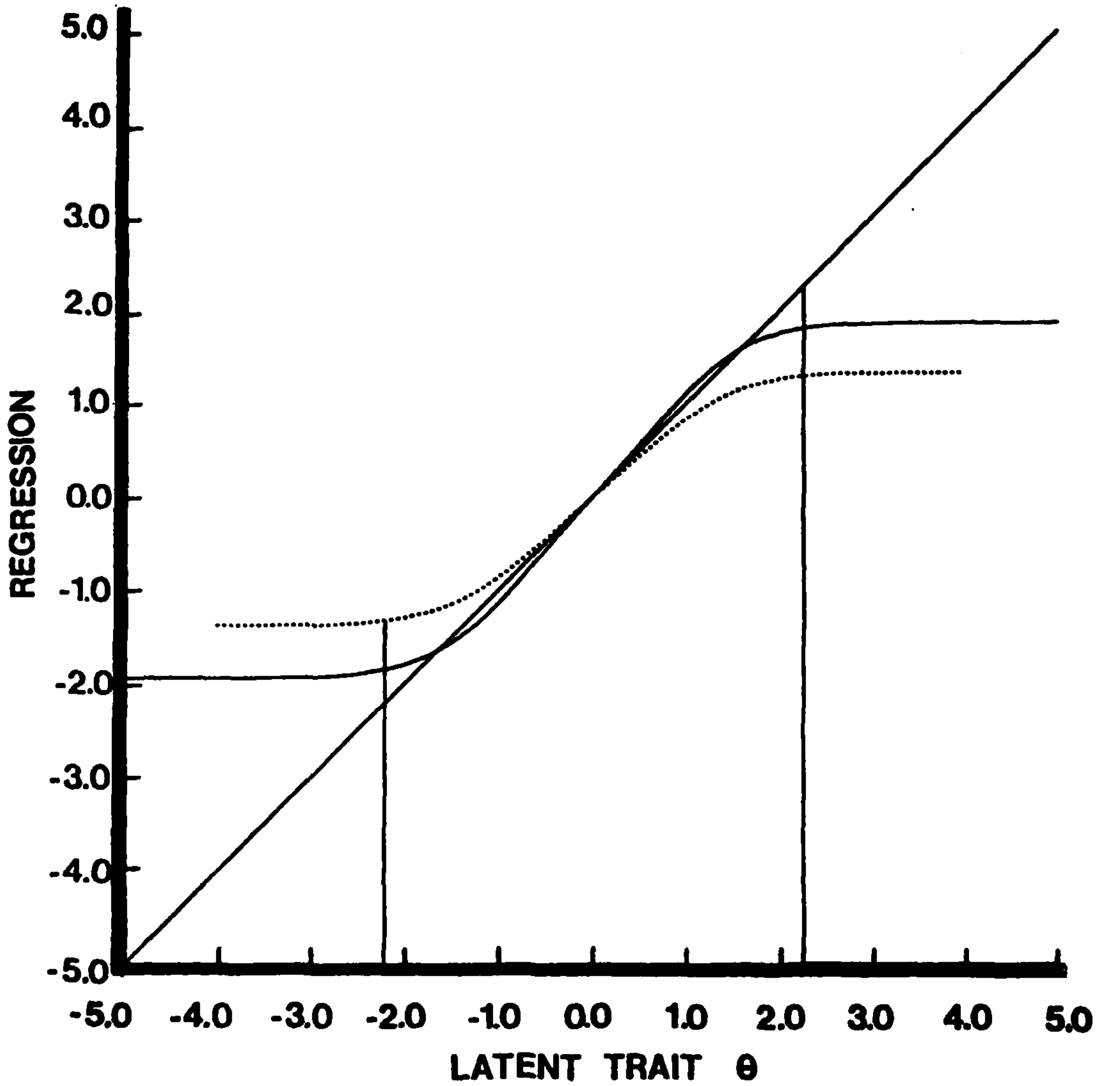

FIGURE 7-1 (Continued)

$\theta^*_{V\text{-min}}$ and $\theta^*_{V\text{-max}}$ Were Obtained by Using $\underline{\theta} = -2.25$ and $\bar{\theta} = 2.25$.

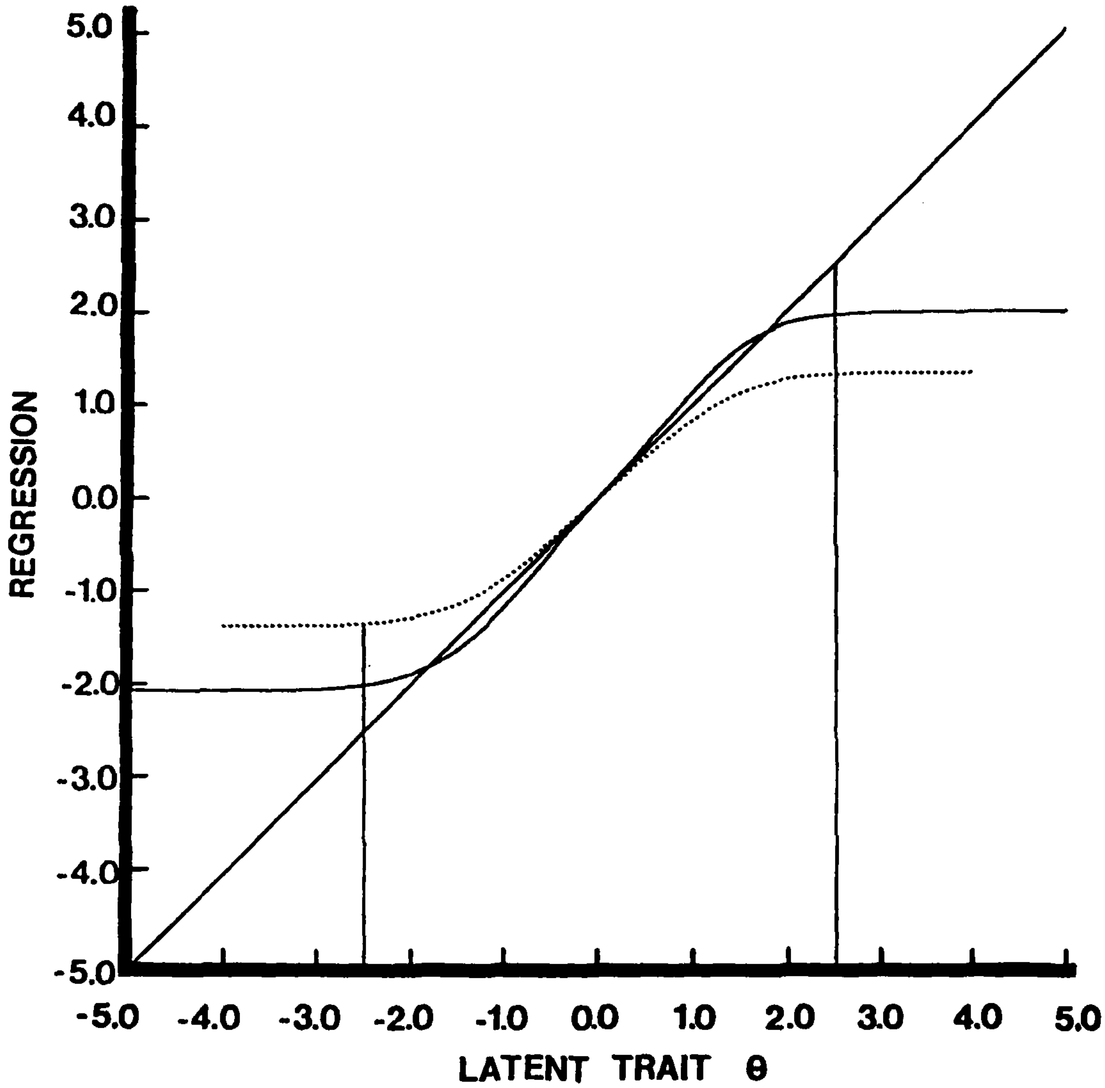

FIGURE 7-1 (Continued)

$\theta^*_{V\text{-}min}$ and $\theta^*_{V\text{-}max}$ Were Obtained by Using $\underline{\theta} = -2.50$ and $\bar{\theta} = 2.50$.

of LIS-U shown in Figure 3-1. As we can see in this figure, the test information function, $I(\theta)$, assumes very small values outside of the interval, $(-2.00,2.00)$, which are, in fact, less than unity, and, therefore, we should not expect it to measure the individual's ability accurately outside this interval. Thus it will be best to use -1.79255 and 1.77649 as the substitute for the maximum likelihood estimates for the two extreme response patterns for this test with the above restriction of the range of θ , or to use either one of the two sets, -1.64702 and 1.65605 with the restricted range of $(-1.75,1.75)$ and -1.47883 and 1.52237 with the restricted range of $(-1.50,1.50)$. These are suitable selections, considering the fact that the least value of the maximum likelihood estimates for the remaining 126 response patterns on LIS-U is -1.3167 for the response pattern, (0,0,0,1,0,0,0) , and the greatest value is 1.3028 for the response pattern, (1,1,1,0,1,1,1) . Similar graphs for the other six cases are presented in Appendix as Figure A-1. We can see that, as the interval, $(\underline{\theta}, \bar{\theta})$, becomes larger, the departure of $E(\theta^*|\theta)$ from the unbiasedness becomes greater, which indicates that these sets of estimates are less and less suitable for use as $\theta^*_{V\text{-min}}$ and $\theta^*_{V\text{-max}}$.

The maximum likelihood estimates for the other 126 response patterns on LIS-U are also presented in Appendix, as Table A-5.

VIII Discussion and Conclusions

Bayesian estimation was considered in comparison with the maximum likelihood estimation from the standpoint of the objectivity of testing, which is closely related with the unbiasedness of estimation and the population-free estimation. Using several different types of tests, including both paper-and-pencil tests and computerized adaptive tests, the effect of priors on the resultant estimate was observed. It was pointed out that the use of priors in Bayesian estimation will result in biases which favor certain individuals over certain other individuals, even though they are exactly equal with respect to their ability levels. An alternative method of using the maximum likelihood estimation with the replacement of positive and negative infinities for the two extreme response patterns, $(0,0,\ldots,0)$ and $(m_1,m_2,\ldots,m_n)$, by a pair of new estimates, $\theta^*_{V\text{-}min}$ and $\theta^*_{V\text{-}max}$, was proposed, and the resulting regression shows less amount of bias than Bayes modal estimate does. We must emphasize that, unlike Bayesian estimates, this modified maximum likelihood estimate, θ^* , is a population-free estimate, so that all the individuals on the same ability level are treated fairly and equally, regardless of the populations from which they are selected.

There exists some philosophical problem in Bayesian estimation which we must call our attention to. In any

Bayesian estimation, we assume the interchangeability of individuals who are assigned to the same prior. The idea of measuring individuals' ability itself preassumes, however, the heterogeneity of individuals, even though they belong to the same population, which implies that the individuals are <u>not</u> <u>interchangeable</u>. In addition to this fact, it should be noted that the assignment of an individual to a specific population is more or less arbitrary. Most researchers use such attributes as age, sex, ethnic background, and so forth, for defining populations. Note that they are only a partial information about an individual, even if we combine a few of these attributes. Thus it happens frequently that a Ph. D. in psychology with a certain ethnic minority background is assigned to the group of Ph. D.'s in psychology, or to the group of people with the same ethnic background. The resultant two Bayesian estimates for this person can be substantially different from each other, depending upon the difference between the two priors. To avoid this contradiction, we must accurately specify the population to which each individual belongs, taking the intersection of thousands of factors, including sex, age, education, ethnic background, etc. If we do this, we will end up with assigning each individual to his own prior, which is shared by no one else. If we know such a prior, however, we do not need to test him at all!

From all aspects, we must conclude that the common belief in the superiority of the Bayesian estimation over the maximum likelihood estimation in the ability measurement is a farce, and the additional information, the prior, is nothing but a resource for the biases, which may lead to unfair personnel selection and other serious social issues.

REFERENCES

[1] Birnbaum, A. Some latent trait models and their use in inferring an examinee's ability. In F. M. Lord & M. R. Novick. (Eds.), Statistical theories of mental test scores. Reading, Mass.: Addison-Wesley, 1968.

[2] Indow, T. & Samejima, F. LIS measurement scale for non-verbal reasoning ability. Tokyo: Nippon Bunka Kagakusha, 1962. (in Japanese)

[3] Indow, T. & Samejima, F. On the results obtained by absolute scaling model and the Lord model in the field of intelligence. Yokohama: Psychological Laboratory, Hiyoshi Campus, Keio University, 1966.

[4] Kendall, M. G. & Stuart, A. The advanced theory of statistics (Vol. 2). New York: Hafner, 1961.

[5] Lord, F. M. and M. R. Novick. Statistical theories of mental test scores. Reading, Mass.: Addison-Wesley, 1968.

[6] Owen, R. J. A Bayesian sequential procedure for quantal response in the context of adaptive mental testing. Journal of the American Statistical Association, 1975, 70, 351-356.

[7] Samejima, F. Estimation of latent ability using a response pattern of graded scores. Psychometrika Monograph, No. 17, 1969.

[8] Samejima, F. A general model for free-response data. Psychometrika Monograph, No. 18, 1972.

[9] Samejima, F. Homogeneous case of the continuous response model. Psychometrika, 1973a, 38, 203-219.

[10] Samejima, F. A comment on Birnbaum's three-parameter logistic model in the latent theory. Psychometrika, 1973b, 38, 221-233.

[11] Samejima, F. Normal ogive model on the continuous response level in the multidimensional latent space. Psychometrika, 1974, 39, 111-121.

[12] Samejima, F. Graded response model of the latent trait theory and tailored testing. Proceedings of the First Conference on Computerized Adaptive Testing, 1975, Civil Service Commission and Office of Naval Research, 1975, pages 5-17.

REFERENCES (Continued)

[13] Samejima, F. Effects of individual optimization in setting the boundaries of dichotomous items on accuracy of estimation. Applied Psychological Measurement, 1977a, 1, 77-94.

[14] Samejima, F. A use of the information function in tailored testing. Applied Psychological Measurement, 1, 1977b, pages 233-247.

[15] Samejima, F. A method of estimating item characteristic functions using the maximum likelihood estimate of ability. Psychometrika, 42, 1977c, pages 163-191.

[16] Samejima, F. Estimation of the operating characteristics of item response categories I: Introduction to the Two-Parameter Beta Method. Office of Naval Research, Research Report 77-1, 1977d.

[17] Samejima, F. Estimation of the operating characteristics of item response categories II: Further development of the Two-Parameter Beta Method. Office of Naval Research, Research Report 78-1, 1978a.

[18] Samejima, F. Estimation of the operating characteristics of item response categories III: The Normal Approach Method and the Pearson System Method. Office of Naval Research, Research Report 78-2, 1978b.

[19] Samejima, F. Estimation of the operating characteristics of item response categories IV: Comparison of the different methods. Office of Naval Research, Research Report 78-3, 1978c.

[20] Samejima, F. Estimation of the operating characteristics of item response categories V: Weighted Sum Procedure in the Conditional P.D.F. Approach. Office of Naval Research, Research Report 78-4, 1978d.

[21] Samejima, F. Estimation of the operating characteristics of item response categories VI: Proportioned Sum Procedure in the Conditional P.D.F. Approach. Office of Naval Research, Research Report 78-5, 1978e.

[22] Samejima, F. Estimation of the operating characteristics of item response categories VII: Bivariate P.D.F. Approach with Normal Approach Method. Office of Naval Research, Research Report 78-6, 1978f.

APPENDIX

TABLE A-1

Sequential Results of Simulated Tailored Testing for Hypothetical Examinees of Eleven Different Ability Levels. Maximum Likelihood Estimation (MLE) Was Used for Ability Estimation.

θ = -2.25

	Item	Score	MLE	Number of Iterations	Information
1/a	5	1			
2	1	1			
3	1	1			
4	1	1			
5	1	1			
6	1	2	-2.7014	5	
7	1	2	-2.3590	4	5.141
8	1	2	-2.1501	3	6.345
9	2	1	-2.2107	3	7.275
10	2	1	-2.2591	3	8.067
11	2	1	-2.2994	3	8.758
12	2	1	-2.3338	3	9.373
13	1	2	-2.2099	4	11.187
14	2	1	-2.2435	3	11.896
15	2	1	-2.2732	3	12.544
16	2	1	-2.2998	3	13.141
17	2	1	-2.3238	3	13.694
18	1	2	-2.2352	4	15.767
19	2	1	-2.2587	3	16.389
20	2	1	-2.2802	3	16.973
21	2	1	-2.3000	3	17.523
22	2	1	-2.3184	3	18.043
23	1	1	-2.3579	3	18.131
24	1	2	-2.2876	4	20.395

TABLE A-1 (Continued)

θ = -1.75

	Item	Score	MLE	Number of Iterations	Information
1	5	1			
2	1	2	-0.8838	5	
3	3	1	-1.5191	5	2.608
4	3	1	-1.6854	4	3.408
5	2	1	-1.9513	4	3.500
6	2	1	-2.1176	4	3.866
7	2	1	-2.2306	4	4.253
8	2	2	-1.8460	5	8.425
9	2	1	-1.9324	4	9.064
10	2	1	-2.0019	4	9.673
11	2	1	-2.0597	4	10.246
12	2	1	-2.1090	3	10.782
13	2	1	-2.1517	4	11.284
14	2	1	-2.1893	3	11.753
15	2	1	-2.2228	3	12.194
16	2	1	-2.2529	4	12.609
17	2	1	-2.2801	3	13.000
18	2	1	-2.3051	3	13.369
19	1	2	-2.2195	4	16.120
20	2	1	-2.2430	3	16.558
21	2	2	-2.1021	4	21.020

θ = -1.25

	Item	Score	MLE	Number of Iterations	Information
1	5	1			
2	1	2	-0.8638	5	
3	3	1	-1.5191	5	2.608
4	3	1	-1.6854	4	3.408
5	2	2	-1.3825	4	6.627
6	3	2	-1.0927	4	9.976
7	3	1	-1.2096	4	12.026
8	3	1	-1.2923	4	13.733
9	3	2	-1.1457	4	17.263
10	3	1	-1.2118	3	19.168
11	3	2	-1.1120	3	22.442

TABLE A-1 (Continued)

θ = -0.75

	Item	Score	MLE	Number of Iterations	Information
1	5	2			
2	9	1	0.9407	5	
3	7	1	0.4811	4	3.119
4	5	1	-0.0375	4	4.695
5	5	1	-0.2520	4	6.053
6	5	1	-0.3811	4	7.144
7	5	1	-0.4714	4	8.048
8	3	1	-0.7827	4	7.260
9	3	2	-0.6865	4	10.269
10	3	2	-0.6208	3	12.720
11	3	2	-0.5709	3	14.810
12	3	2	-0.5307	3	16.642
13	3	2	-0.4971	3	18.278
14	3	1	-0.6369	3	21.053

θ = -0.25

	Item	Score	MLE	Number of Iterations	Information
1	5	1			
2	1	2	-0.8838	5	
3	3	2	-0.4710	4	3.503
4	3	2	-0.3430	4	4.635
5	5	1	-0.4870	4	6.796
6	3	2	-0.4081	3	8.111
7	5	2	-0.1308	4	8.579
8	5	1	-0.2515	3	11.055
9	5	2	-0.0905	4	12.424
10	5	1	-0.1794	3	14.751
11	5	1	-0.2471	3	16.796
12	5	1	-0.3015	3	18.627
13	5	1	-0.3469	3	20.288

TABLE A-1 (Continued)

θ = 0.00

	Item	Score	MLE	Number of Iterations	Information
1	5	1			
2	1	2	-0.8838	5	
3	3	2	-0.4710	4	3.503
4	3	2	-0.3430	4	4.635
5	5	2	0.0493	4	5.076
6	5	2	0.2535	4	6.172
7	5	1	0.0270	4	9.283
8	5	2	0.1536	4	10.686
9	5	2	0.2466	4	11.966
10	5	1	0.1110	4	15.001
11	5	1	0.0142	3	17.583
12	5	2	0.0870	3	19.241
13	5	1	0.0115	3	21.718

θ = 0.50

	Item	Score	MLE	Number of Iterations	Information
1	5	2			
2	9	1	0.9407	5	
3	7	1	0.4811	4	3.119
4	5	2	0.6395	4	4.387
5	7	1	0.4899	3	6.060
6	5	2	0.5839	3	7.395
7	7	2	0.8538	4	8.352
8	7	1	0.7390	3	10.365
9	7	1	0.6570	3	12.067
10	7	1	0.5935	3	13.544
11	7	1	0.5420	3	14.851
12	7	1	0.4987	3	16.026
13	5	2	0.5354	3	17.739
14	7	2	0.6605	3	19.795
15	7	2	0.7584	3	21.734

TABLE A-1 (Continued)

$\theta = 1.00$

	Item	Score	MLE	Number of Iterations	Information
1	5	2			
2	9	1	0.9407	5	
3	7	2	1.5104	4	2.664
4	8	1	1.1662	4	4.921
5	8	2	1.5034	3	7.267
6	8	1	1.3363	3	9.476
7	8	1	1.2298	3	11.245
8	8	2	1.3913	3	14.089
9	8	1	1.3099	3	15.589
10	8	2	1.4186	3	18.699
11	8	1	1.3531	3	20.678
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]

$\theta = 1.50$

	Item	Score	MLE	Number of Iterations	Information
1	5	2			
2	9	1	0.9407	5	
3	7	2	1.5104	4	2.664
4	8	1	1.1662	4	4.921
5	8	1	1.0231	4	6.407
6	7	2	1.1953	3	8.753
7	8	2	1.4132	3	11.117
8	8	2	1.5485	3	13.121
9	8	1	1.4398	3	15.665
10	8	1	1.3593	3	17.841
11	8	1	1.2959	3	19.757
12	8	1	1.2438	3	21.476

TABLE A-1 (Continued)

$\theta = 2.00$

	Item	Score	MLE	Number of Iterations	Information
1	5	2			
2	9	2			
3	9	2			
4	9	2			
5	9	2			
6	9	1	2.5611	5	
7	9	1	2.2872	4	8.033
8	9	1	2.1201	3	9.913
9	9	1	2.0001	3	11.467
10	8	1	1.8017	3	13.157
11	8	2	1.8583	3	15.166
12	8	1	1.7274	3	17.256
13	8	2	1.7783	3	19.332
14	8	2	1.8208	3	21.210

$\theta = 2.50$

	Item	Score	MLE	Number of Iterations	Information
1	5	2			
2	9	2			
3	9	1	2.0004	4	
4	8	2	2.1419	4	4.134
5	9	2	2.3430	4	4.757
6	9	2	2.4785	4	5.334
7	9	2	2.5784	5	5.855
8	9	2	2.6563	5	6.325
9	9	2	2.7197	4	6.748
10	9	2	2.7728	4	7.133
11	9	2	2.8183	3	7.485
12	9	2	2.8579	4	7.809
13	9	2	2.8930	3	8.109
14	9	2	2.9244	3	8.388
15	9	2	2.9527	3	8.648
16	9	2	2.9785	3	8.893
17	9	1	2.7440	4	13.715
18	9	1	2.5966	4	17.441
19	9	2	2.6237	3	17.965
20	9	2	2.6489	3	18.467
21	9	1	2.5415	4	21.818

TABLE A-2

Sequential Results of Simulated Tailored Testing for Hypothetical Examinees of Eleven Different Ability Levels. Bayes Modal Estimation (BME) Was Used for Ability Estimation, with the Prior, n(0.0,1.0) .

$\theta = -2.25$

	Item	Score	BME	Number of Iterations	Information
1	**	*	0.0	**	
2	5	1	-0.5385	5	1.454
3	3	1	-1.1984	5	2.726
4	3	1	-1.4223	4	4.053
5	3	1	-1.5456	4	4.977
6	3	1	-1.6282	4	5.670
7	2	1	-1.7953	4	5.373
8	2	1	-1.9286	4	5.394
9	2	1	-2.0348	4	5.596
10	2	1	-2.1205	4	5.878
11	2	1	-2.1909	4	6.185
12	2	1	-2.2497	4	6.490
13	2	2	-1.9812	5	11.409
14	2	1	-2.0303	4	11.887
15	2	1	-2.0735	4	12.356
16	2	1	-2.1121	3	12.812
17	2	1	-2.1467	3	13.252
18	2	1	-2.1781	4	13.675
19	2	1	-2.2067	3	14.080
20	2	1	-2.2329	3	14.469
21	2	1	-2.2571	3	14.842
22	2	1	-2.2796	4	15.199
23	2	1	-2.3004	3	15.542
24	2	1	-2.3199	3	15.871
25	1	2	-2.2479	5	18.809
26	2	1	-2.2665	3	19.194
27	2	1	-2.2840	3	19.565
28	2	1	-2.3005	3	19.922
29	2	1	-2.3161	3	20.268

TABLE A-2 (Continued)

θ = -1.75

	Item	Score	BME	Number of Iterations	Information
1	**	*	0.0	**	
2	5	1	-0.5385	5	1.454
3	3	1	-1.1984	5	2.726
4	3	1	-1.4223	4	4.053
5	3	1	-1.5456	4	4.977
6	3	1	-1.6282	4	5.670
7	2	1	-1.7953	4	5.373
8	12	2	-1.5831	4	9.421
9	2	1	-1.6928	4	9.673
10	2	2	-1.5620	4	12.918
11	2	1	-1.6458	4	13.461
12	2	1	-1.7174	4	14.040
13	2	2	-1.6179	4	17.049
14	2	2	-1.5413	3	19.673
15	3	1	-1.5660	3	20.936

θ = -1.25

	Item	Score	BME	Number of Iterations	Information
1	**	*	0.0	**	
2	5	1	-0.5385	5	1.454
3	3	2	-0.3693	3	3.150
4	5	1	-0.5195	3	4.788
5	3	4	-0.4268	3	6.429
6	5	1	-0.5141	3	8.087
7	3	2	-0.4499	3	9.713
8	3	2	-0.4019	3	11.049
9	5	1	-0.4624	3	12.998
10	3	1	-0.6683	3	15.617
11	3	1	-0.8044	3	18.157
12	3	2	-0.7495	3	20.398

TABLE A-2 (Continued)

$\theta = -0.75$

	Item	Score	BME	Number of Iterations	Information
1	**	*	0.0	**	
2	5	1	-0.5385	5	1.454
3	3	2	-0.3693	3	3.150
4	5	1	-0.5195	3	4.788
5	3	2	-0.4268	3	6.429
6	5	1	-0.5141	3	8.087
7	3	2	-0.4499	3	9.713
8	3	2	-0.4019	3	11.049
9	5	1	-0.4624	3	12.998
10	3	2	-0.4233	3	14.388
11	5	2	-0.2654	4	15.069
12	5	1	-0.3218	3	17.277
13	5	2	-0.2073	3	18.564
14	5	1	-0.2581	3	20.736

$\theta = -0.25$

	Item	Score	BME	Number of Iterations	Information
1	**	*	0.0	**	
2	5	1	-0.5385	5	1.454
3	3	2	-0.3693	3	3.150
4	5	1	-0.5195	3	4.788
5	3	2	-0.4268	3	6.429
6	5	1	-0.5141	3	8.087
7	3	2	-0.4499	3	9.713
8	3	2	-0.4019	3	11.049
9	5	1	-0.4624	3	12.998
10	3	2	-0.4233	3	14.388
11	5	1	-0.4703	3	16.283
12	3	2	-0.4172	3	17.709
13	5	1	-0.4756	3	19.568
14	3	2	-0.4471	3	21.019

TABLE A-2 (Continued)

θ = 0.00

	Item	Score	BME	Number of Iterations	Information
1	**	*	0.0	**	
2	5	1	-0.5385	5	1.454
3	3	2	-0.3693	3	3.150
4	5	1	-0.5195	3	4.788
5	3	2	-0.4268	3	6.429
6	5	1	-0.5141	3	8.087
7	3	2	-0.4499	3	9.713
8	3	2	-0.4019	3	11.049
9	5	1	-0.4624	3	12.998
10	3	2	-0.4233	3	14.388
11	5	2	-0.2654	4	15.069
12	5	2	-0.1448	4	16.168
14	5	1	-0.2073	3	18.564
14	5	2	-0.1156	3	20.011

θ = 0.50

	Item	Score	BME	Number of Iterations	Information
1	**	*	0.0	**	
2	5	2	0.5385	5	1.454
3	7	2	1.1527	5	2.154
4	8	1	0.9834	3	4.054
5	7	2	1.1988	4	5.890
6	8	1	1.0871	3	7.769
7	8	1	1.0121	3	9.255
8	7	2	1.1302	3	11.572
9	8	1	1.0698	3	13.153
10	7	2	1.1537	3	15.392
11	8	1	1.1036	3	17.034
12	8	1	1.0620	3	18.494
13	7	1	0.9854	3	19.464
14	7	1	0.9204	3	20.437

TABLE A-2 (Continued)

θ = 1.00

	Item	Score	BME	Number of Iterations	Information
1	**	*	0.0	**	
2	5	2	0.5385	5	1.454
3	7	2	1.1597	5	2.156
4	8	1	0.9834	3	4.054
5	7	1	0.7788	3	5.597
6	7	1	0.6545	3	6.938
7	7	1	0.5666	3	8.100
8	7	1	0.4993	3	9.121
9	5	2	0.5518	3	10.931
10	7	1	0.4993	3	12.056
11	5	2	0.5423	3	13.800
12	7	1	0.4993	3	14.952
13	5	2	0.5357	3	16.690
14	7	1	0.4993	3	17.928
15	5	2	0.5308	3	19.592
16	5	2	0.5583	3	21.111

θ = 1.50

	Item	Score	BME	Number of Iterations	Information
1	**	*	0.0	**	
2	5	2	0.5385	5	1.454
3	7	2	1.1597	5	2.156
4	8	2	1.6435	5	3.465
5	8	1	1.3725	4	6.259
6	8	2	1.5688	4	8.226
7	8	2	1.6948	4	9.875
8	8	1	1.5455	4	12.880
9	8	1	1.4447	3	15.353
10	8	1	1.3689	3	17.495
11	8	2	1.4569	3	19.936
12	8	2	1.5272	3	22.120

TABLE A-2 (Continued)

$\theta = 2.00$

Item	Score		BME	Number of Iterations	Information
1	**	*	0.0	**	
2	5	2	0.5385	5	1.454
3	7	2	1.1597	5	2.156
4	8	1	0.9834	3	4.054
5	7	2	1.1988	4	5.890
6	8	2	1.4628	4	7.635
7	8	2	1.6202	4	9.253
8	8	1	1.4762	4	12.187
9	8	2	1.5821	3	14.016
10	8	2	1.6631	3	15.648
11	8	1	1.5624	3	18.702
12	8	2	1.6272	3	20.462

$\theta = 2.50$

Item	Score		BME	Number of Iterations	Information
1	**	*	0.0	**	
2	5	2	0.5385	5	1.454
3	7	2	1.1597	5	2.156
4	8	2	1.6435	5	3.465
5	8	2	1.8574	4	4.714
6	8	2	1.9853	4	5.658
7	8	2	2.0739	4	6.397
8	8	2	2.1407	4	6.999
9	9	1	1.9883	4	10.405
10	8	2	2.0422	4	11.259
11	8	2	2.0870	4	12.007
12	8	2	2.1251	4	12.672
13	9	2	2.2060	4	12.409
14	9	2	2.2788	4	12.271
15	9	2	2.3444	4	12.244
16	9	1	2.2248	4	16.102
17	9	2	2.2809	4	16.256
18	9	2	2.3303	4	16.464
19	9	1	2.2392	4	19.929
20	9	2	2.2822	3	20.253

TABLE A-3

Sequential Results of Simulated Tailored Testing for Hypothetical Examinees of Eleven Different Ability Levels. Bayes Modal Estimation (BME) Was Used for Ability Estimation, with the Prior, n(0.0,0.8) .

$\theta = -2.25$

	Item	Score	BME	Number of Iterations	Information
1	**	*	0.0	**	
2	5	1	-0.4334	4	1.648
3	5	1	-0.5986	4	2.673
4	3	1	-1.0601	5	3.501
5	3	1	-1.2829	4	5.010
6	3	1	-1.4113	5	6.244
7	3	1	-1.4993	4	7.230
8	3	1	-1.5652	4	8.039
9	2	1	-1.6946	4	7.768
10	2	1	-1.8037	4	7.741
11	2	1	-1.8959	5	7.890
12	2	1	-1.9741	4	8.148
13	2	2	-1.7915	5	12.431
14	2	1	-1.8525	4	12.866
15	2	1	-1.9062	4	13.328
16	2	1	-1.9538	3	13.808
17	2	1	-1.9904	3	14.291
18	2	1	-2.0347	3	14.770
19	2	1	-2.0694	3	15.242
20	2	1	-2.1011	3	15.702
21	2	1	-2.1301	3	16.149
22	2	1	-2.1568	4	16.583
23	2	1	-2.1810	3	17.002
24	2	1	-2.2046	3	17.408
25	2	1	-2.2260	4	17.800
26	2	1	-2.2461	3	18.178
27	2	1	-2.2650	3	18.544
28	2	1	-2.2826	3	18.697
29	2	1	-2.2995	3	19.239
30	2	1	-2.3154	3	19.565
31	1	1	-2.3502	5	19.249
32	1	1	-2.3839	4	18.972
33	1	1	-2.4164	4	18.734
34	1	1	-2.4479	4	18.535
35	1	1	-2.4704	3	18.372
36	1	2	-2.4057	4	21.475

TABLE A-3 (Continued)

$\theta = -1.75$

	Item	Score	BME	Number of Iterations	Information
1	**	*	0.0	**	
2	5	1	-0.4334	4	1.648
3	5	1	-0.5986	4	2.673
4	3	1	-1.0661	5	3.501
5	3	1	-1.2829	4	5.010
6	3	1	-1.4113	5	6.244
7	3	1	-1.4993	4	7.230
8	3	1	-1.5652	4	8.039
9	2	1	-1.6946	4	7.768
10	2	2	-1.5470	4	11.554
11	3	1	-1.5864	4	12.487
12	2	2	-1.4916	4	15.762
13	3	1	-1.5264	3	16.850
14	3	1	-1.5568	3	17.836
15	2	2	-1.4873	4	21.041

$\theta = -1.25$

	Item	Score	BME	Number of Iterations	Information
1	**	*	0.0	**	
2	5	1	-0.4334	4	1.648
3	5	1	-0.5986	4	2.673
4	3	1	-1.0661	5	3.501
5	3	2	-0.8387	4	6.627
6	3	2	-0.7217	3	8.999
7	3	2	-0.6433	3	10.939
8	3	1	-0.8133	3	13.921
9	3	2	-0.7473	3	15.993
10	3	1	-0.8597	3	18.967
11	3	1	-0.9449	3	21.627

TABLE A-3 (Continued)

θ = -0.75

Item	Score	BME	Number of Iterations	Information
1	**	*	0.0	**	
2	5	1	-0.4334	4	1.648
3	5	1	-0.5986	4	2.673
4	3	2	-0.4677	3	4.841
5	3	2	-0.3925	3	6.354
6	5	1	-0.4797	3	8.095
7	3	2	-0.4240	3	9.634
8	5	1	-0.4858	3	11.364
9	3	1	-0.7022	3	13.457
10	3	1	-0.8448	3	15.694
11	3	2	-0.7807	3	18.133
12	3	2	-0.7296	3	20.309

θ = -0.25

Item	Score	BME	Number of Iterations	Information
1	**	*	0.0	**	
2	5	1	-0.4334	4	1.648
3	5	2	0.0000	3	4.125
4	5	2	0.1887	4	5.933
5	5	1	-0.0000	3	8.251
6	5	2	0.1219	3	10.134
7	5	1	-0.0000	3	12.376
8	5	1	-0.0901	3	14.301
9	5	1	-0.1612	3	16.003
10	5	2	-0.0716	3	18.452
11	5	1	-0.1306	3	20.216

TABLE A-3 (Continued)

$\theta = 0.00$

	Item	Score	BME	Number of Iterations	Information
1	**	*	0.0	**	
2	5	1	-0.4334	4	1.648
3	5	1	-0.5986	4	2.673
4	3	2	-0.4677	3	4.841
5	3	2	-0.3925	3	6.354
6	5	2	-0.1384	4	7.658
7	5	2	0.0207	4	9.224
8	5	1	-0.0945	3	11.621
9	5	1	-0.1782	3	13.715
10	5	2	-0.0716	3	15.673
11	5	2	0.0117	3	17.507
12	5	1	-0.0576	3	19.756
13	5	2	0.0096	3	21.640

$\theta = 0.50$

	Item	Score	BME	Number of Iterations	Information
1	**	*	0.0	**	
2	5	2	0.4334	4	1.648
3	5	2	0.5986	4	2.673
4	7	1	0.4636	4	4.540
5	5	2	0.5598	4	5.735
6	7	1	0.4748	3	7.460
7	5	2	0.5429	3	8.728
8	7	2	0.7576	4	9.260
9	7	1	0.6780	3	11.302
10	7	2	0.8226	4	12.468
11	7	1	0.7539	3	14.457
12	7	2	0.8607	3	15.902
13	7	1	0.8015	3	17.850
14	7	1	0.7521	3	19.635
15	7	1	0.7092	3	21.284

TABLE A-3 (Continued)

θ = 1.00

	Item	Score	BME	Number of Iterations	Information
1	**	*	0.0	**	
2	5	2	0.4334	4	1.648
3	5	2	0.5986	4	2.673
4	7	1	0.4636	4	4.540
5	5	2	0.5598	4	5.735
6	7	1	0.4748	3	7.460
7	5	2	0.5429	3	8.728
8	7	1	0.4807	3	10.388
9	5	2	0.5334	3	11.695
10	5	2	0.5767	3	12.839
11	7	1	0.5274	3	14.652
12	5	2	0.5638	3	15.844
13	7	1	0.5231	3	17.602
14	5	2	0.5547	3	18.830
15	7	1	0.5200	3	20.549

θ = 1.50

	Item	Score	BME	Number of Iterations	Information
1	**	*	0.0	**	
2	5	2	0.4334	4	1.648
3	5	2	0.5986	4	2.673
4	7	2	1.0203	5	2.943
5	7	2	1.2553	4	3.966
6	8	2	1.5617	4	5.063
7	8	2	1.7400	4	6.350
8	8	1	1.5354	4	9.763
9	8	1	1.4163	3	12.458
10	8	1	1.3325	3	14.749
11	8	2	1.4375	3	16.985
12	8	2	1.5192	3	19.024
13	8	1	1.4501	3	21.542

TABLE A-3 (Continued)

θ = 2.00

Item	Score		BME	Number of Iterations	Information
1	**	*	0.0	**	
2	5	2	0.4334	4	1.648
3	5	2	0.5986	4	2.673
4	7	2	1.0203	5	2.943
5	7	2	1.2553	4	3.966
6	8	2	1.5617	4	5.063
7	8	1	1.3742	4	8.015
8	8	2	1.5354	4	9.763
9	8	2	1.6484	4	11.368
10	8	1	1.5249	4	14.403
11	8	2	1.6084	3	16.171
12	8	2	1.6756	3	17.781
13	8	1	1.5856	4	20.892

θ = 2.50

Item	Score		BME	Number of Iterations	Information
1	**	*	0.0	**	
2	5	2	0.4334	4	1.648
3	5	2	0.5986	4	2.673
4	7	2	1.0203	5	2.943
5	7	2	1.2553	4	3.966
6	8	2	1.5617	4	5.063
7	8	2	1.7400	4	6.350
8	8	2	1.8589	4	7.476
9	8	2	1.9457	4	8.431
10	8	2	2.0130	4	9.247
11	8	2	2.0675	4	9.955
12	8	2	2.1140	4	10.577
13	9	2	2.2072	4	10.192
14	9	1	2.0891	4	13.878
15	8	2	2.1209	3	14.618
16	9	2	2.1893	5	14.512
17	9	1	2.1010	4	17.895
18	9	2	2.1585	4	18.004
19	9	2	2.2111	4	18.173
20	9	2	2.2593	4	18.394
21	9	2	2.3037	3	18.658
22	9	1	2.2251	4	21.966

TABLE A-4

Sequential Results of Simulated Tailored Testing for Hypothetical Examinees of Eleven Different Ability Levels. Bayes Modal Estimation (BME) Was Used for Ability Estimation, with the Prior, n(0.0,0.5) .

θ = -2.25

	Item	Score	BME	Number of Iterations	Information
1	**	*	0.0	**	
2	5	1	-0.2436	4	1.923
3	5	1	-0.3785	4	3.478
4	5	1	-0.4700	4	4.748
5	3	1	-0.7746	4	5.318
6	3	1	-0.9658	4	6.997
7	3	1	-1.0968	4	8.866
8	3	1	-1.1931	5	10.630
9	3	1	-1.2678	4	12.225
10	3	1	-1.3281	4	13.655
11	3	1	-1.3781	4	14.941
12	3	1	-1.4207	4	16.104
13	3	1	-1.4576	3	17.161
14	3	1	-1.4900	3	18.130
15	3	1	-1.5189	3	19.022
16	3	1	-1.5448	3	15.848
17	3	1	-1.5683	3	20.617

TABLE A-4 (Continued)

$\theta = -1.75$

	Item	Score	BME	Number of Iterations	Information
1	**	*	0.0	**	
2	5	1	-0.2436	4	1.923
3	5	1	-0.3785	4	3.478
4	5	1	-0.4700	4	4.748
5	3	1	-0.7746	4	5.318
6	3	1	-0.9658	4	6.557
7	3	1	-1.0968	4	8.866
8	3	1	-1.1931	5	10.630
9	3	1	-1.2678	4	12.225
10	3	1	-1.3281	4	13.655
11	3	1	-1.3781	4	14.941
12	3	1	-1.4207	4	16.104
13	3	1	-1.4576	3	17.161
14	3	1	-1.4900	3	18.130
15	3	1	-1.5189	3	19.022
16	3	1	-1.5448	3	19.848
17	3	1	-1.5683	3	20.617

$\theta = -1.25$

	Item	Score	BME	Number of Iterations	Information
1	**	*	0.0	**	
2	5	1	-0.2436	4	1.923
3	5	1	-0.3785	4	3.478
4	5	1	-0.4700	4	4.748
5	3	1	-0.7746	4	5.318
6	3	2	-0.6830	3	7.895
7	3	1	-0.8551	4	9.903
8	3	1	-0.9768	4	12.037
9	3	2	-0.8941	3	14.807
10	3	2	-0.8312	3	17.285
11	3	1	-0.9167	3	19.808
12	3	1	-0.9859	3	22.182

TABLE A-4 (Continued)

θ = -0.75

Item	Score	BME	Number of Iterations	Information
1	**	*	0.0	**	
2	5	2	0.2436	4	1.923
3	5	1	0.0000	3	4.125
4	5	1	-0.1424	3	6.042
5	5	1	-0.2414	3	7.701
6	5	1	-0.3165	3	9.157
7	5	1	-0.3766	3	10.452
8	5	1	-0.4264	3	11.618
9	5	1	-0.4688	3	12.680
10	3	1	-0.6425	4	12.097
11	3	2	-0.6000	3	14.688
12	3	2	-0.5653	3	16.991
13	3	2	-0.5359	3	19.071
14	3	2	-0.5104	3	20.975

θ = -0.25

Item	Score	BME	Number of Iterations	Information
1	**	*	0.0	**	
2	5	2	0.2436	4	1.923
3	5	1	0.0000	3	4.125
4	5	1	-0.1424	3	6.042
5	5	1	-0.2414	3	7.701
6	5	1	-0.3165	3	9.157
7	5	1	-0.3766	3	10.452
8	5	2	-0.2407	4	13.482
9	5	1	-0.2928	3	14.906
10	5	2	-0.1947	3	17.751
11	5	1	-0.2404	3	19.264
12	5	1	-0.2802	3	20.673

TABLE A-4 (Continued)

$\theta = 0.00$

	Item	Score	BME	Number of Iterations	Information
1	**	*	0.0	**	
2	5	1	-0.2436	4	1.923
3	5	2	0.0000	3	4.125
4	5	2	0.1424	3	6.042
5	5	1	0.0000	3	8.251
6	5	2	0.1009	3	10.190
7	5	2	0.1783	3	11.920
8	5	1	0.0781	3	14.335
9	5	2	0.1416	3	16.116
10	5	2	0.1947	3	17.751
11	5	1	0.1175	3	20.294

$\theta = 0.50$

	Item	Score	BME	Number of Iterations	Information
1	**	*	0.0	**	
2	5	1	-0.2436	4	1.923
3	5	2	0.0000	3	4.125
4	5	1	-0.1424	3	6.042
5	5	2	0.0000	3	8.251
6	5	2	0.1009	3	10.190
7	5	2	0.1783	3	11.920
8	5	2	0.2407	3	13.482
9	5	2	0.2928	3	14.906
10	5	2	0.3373	3	16.216
11	5	1	0.2404	3	19.264
12	5	2	0.2802	3	20.673

TABLE A-4 (Continued)

θ = 1.00

Item	Score		BME	Number of Iterations	Information
1	**	*	0.0	**	
2	5	2	0.2436	4	1.923
3	5	2	0.3785	4	3.478
4	5	2	0.4700	4	4.748
5	5	2	0.5385	4	5.817
6	7	2	0.7620	4	5.771
7	7	2	0.9298	4	6.452
8	7	1	0.8298	4	8.868
9	7	2	0.9485	4	10.004
10	7	1	0.8684	3	12.260
11	7	2	0.9594	3	13.612
12	7	2	1.0352	3	14.981
13	7	2	1.0995	3	16.326
14	8	2	1.2152	4	17.300
15	8	1	1.1755	3	19.636
16	8	2	1.2660	3	21.096

θ = 1.50

Item	Score		BME	Number of Iterations	Information
1	**	*	0.0	**	
2	5	2	0.2436	4	1.923
3	5	2	0.3785	4	3.478
4	5	2	0.4700	4	4.748
5	5	2	0.5385	4	5.817
6	7	2	0.7620	4	5.771
7	7	1	0.6777	4	8.007
8	7	1	0.6150	3	9.930
9	7	1	0.5651	3	11.627
10	7	1	0.5236	3	13.151
11	5	2	0.5556	3	14.559
12	7	2	0.6730	3	15.793
13	7	2	0.7691	3	17.152
14	7	1	0.7260	3	19.000
15	7	2	0.8041	3	20.540

TABLE A-4 (Continued)

θ = 2.00

	Item	Score	BME	Number of Iterations	Information
1	**	*	0.0	**	
2	5	2	0.2436	4	1.923
3	5	2	0.3785	4	3.478
4	5	2	0.4700	4	4.748
5	5	2	0.5385	4	5.817
6	7	2	0.7620	4	5.771
7	7	2	0.9298	4	6.452
8	7	2	1.0584	4	7.484
9	7	2	1.1597	4	8.626
10	8	2	1.3319	5	9.567
11	8	2	1.4615	4	10.931
12	8	2	1.5624	4	12.407
13	8	2	1.6435	4	13.862
14	8	2	1.7103	4	15.243
15	8	2	1.7667	4	16.536
16	8	2	1.8152	3	17.739
17	8	2	1.8574	4	18.858
18	8	1	1.7564	4	23.023

θ = 2.50

	Item	Score	BME	Number of Iterations	Information
1	**	*	0.0	**	
2	5	2	0.2436	4	1.923
3	5	2	0.3785	4	3.478
4	5	2	0.4700	4	4.748
5	5	2	0.5385	4	5.817
6	7	2	0.7620	4	5.771
7	7	2	0.9298	4	6.452
8	7	2	1.0584	4	7.484
9	7	2	1.1597	4	8.626
10	8	2	1.3319	5	9.567
11	8	2	1.4615	4	10.931
12	8	2	1.5624	4	12.407
13	8	2	1.6435	4	13.862
14	8	2	1.7103	4	15.243
15	8	2	1.7667	4	16.536
16	8	2	1.8152	3	17.739
17	8	2	1.8574	4	18.858
18	8	2	1.8947	3	19.899
19	8	2	1.9280	3	20.871

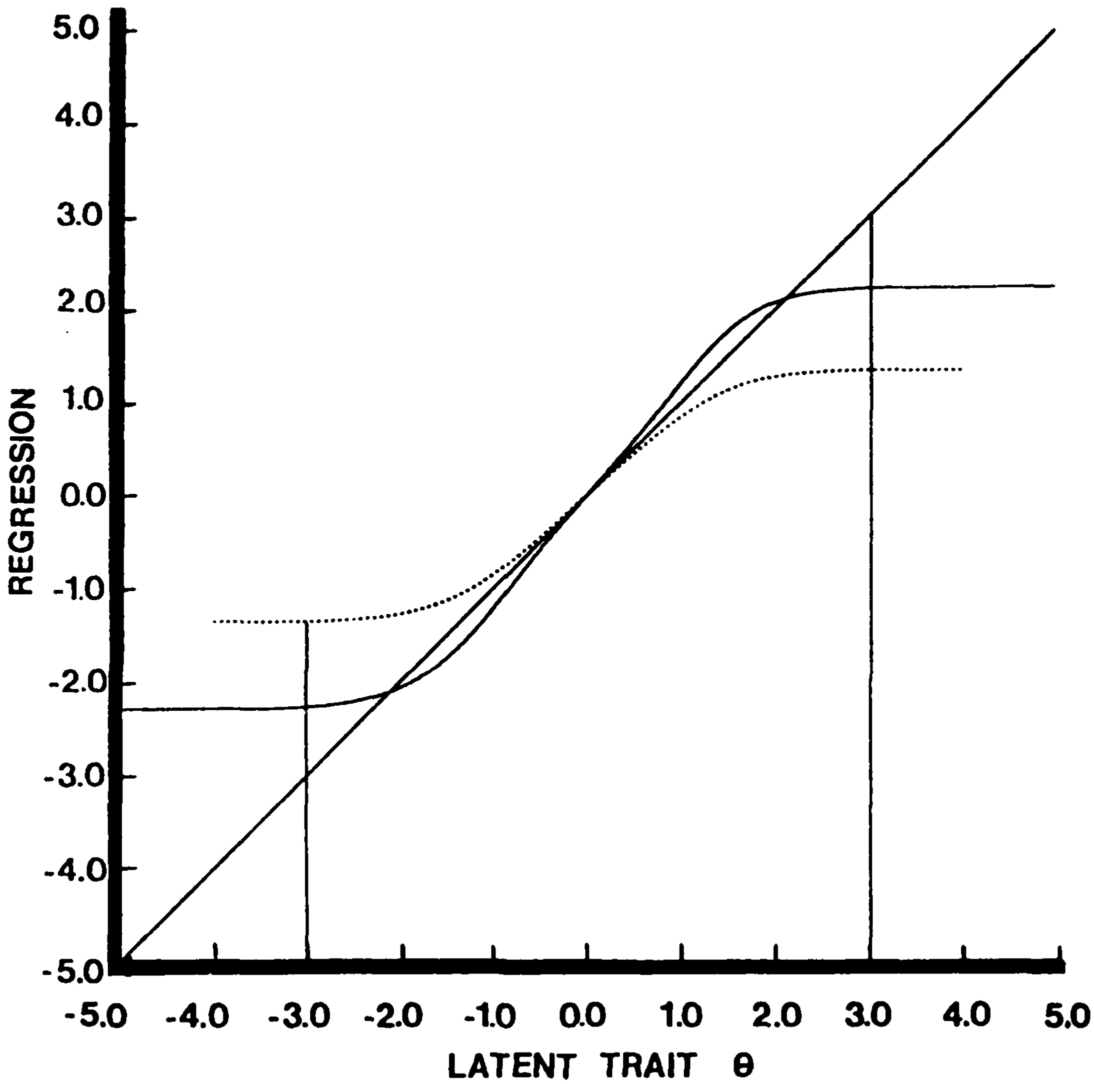

FIGURE A-1

Regression of the Maximum Likelihood Estimate with Those for the Two Extreme Response Patterns, (0,0,...,0) and (1,1,...,1), Replaced by $\theta^*_{V\text{-}min}$ and $\theta^*_{V\text{-}max}$, respectively, on Ability θ (Solid Curve), Together with the Regression of the Bayes Modal Estimate with n(0,1) as the Prior (Dotted Curve). These Two Estimates Were Obtained by Using $\underline{\theta} = -3.00$ and $\bar{\theta} = 3.00$.

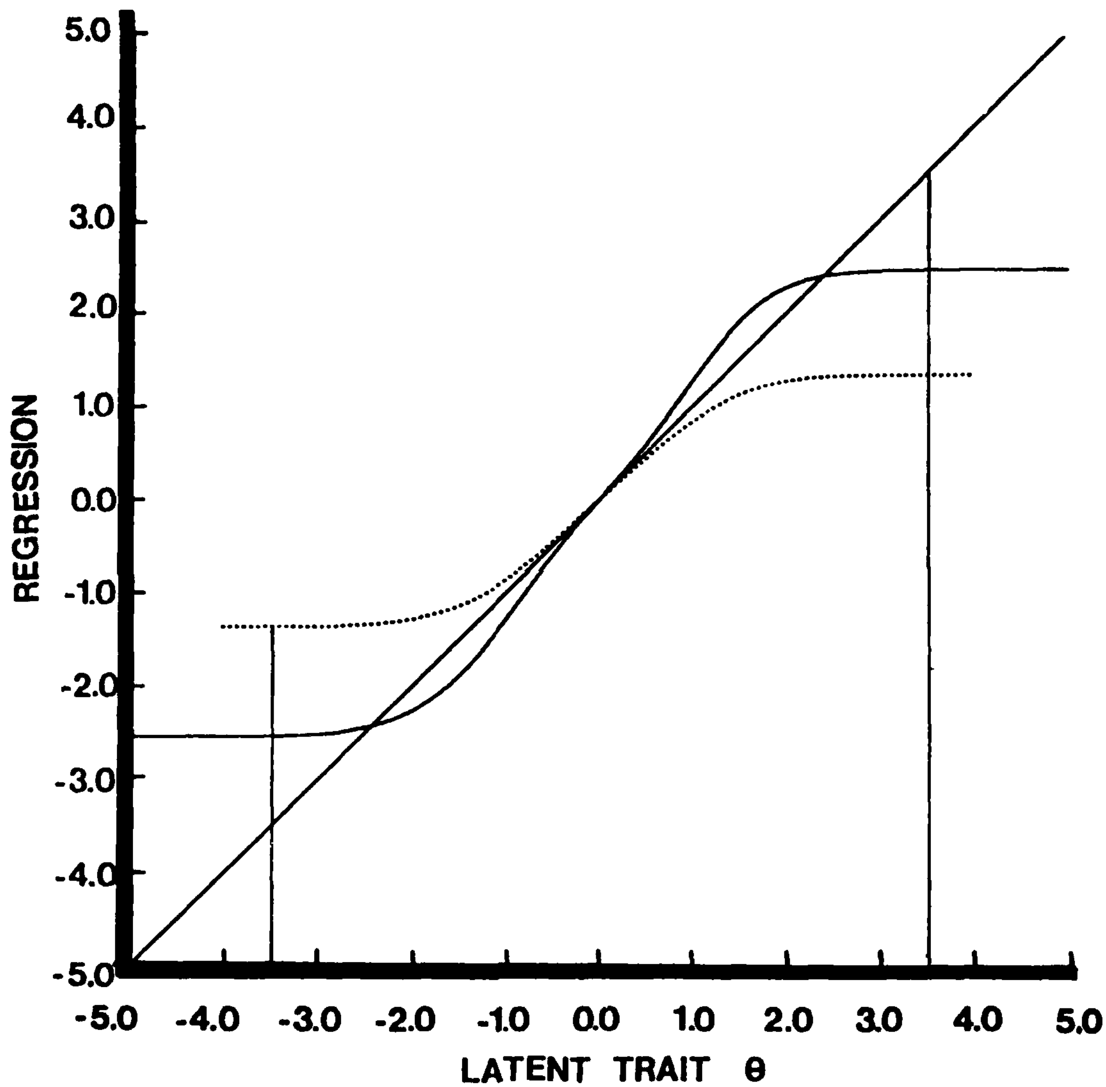

FIGURE A-1 (Continued)

θ^*_{V-min} and θ^*_{V-max} Were Obtained by Using $\underline{\theta} = -3.50$ and $\bar{\theta} = 3.50$.

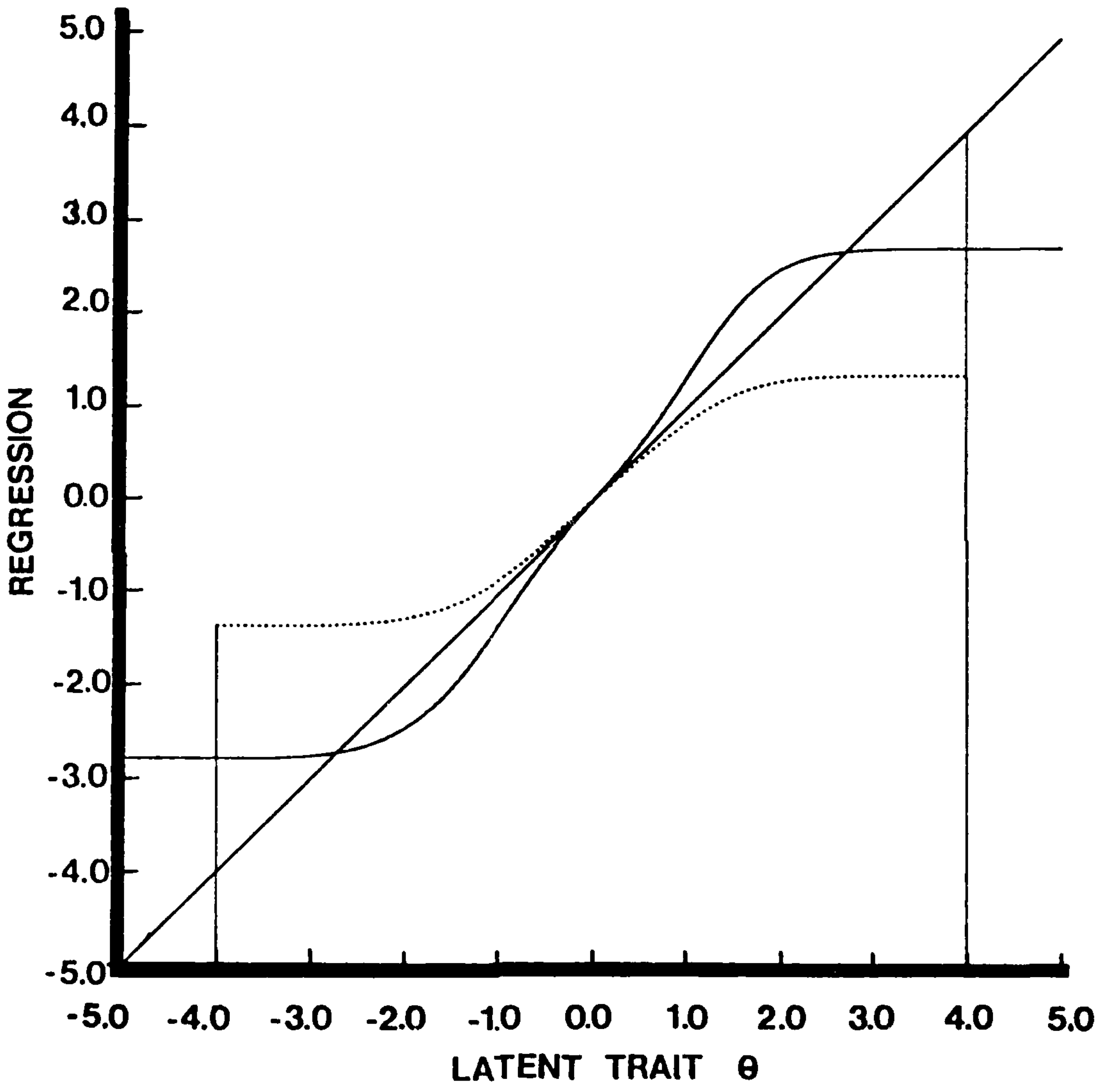

FIGURE A-1 (Continued)

$\theta^*_{V\text{-min}}$ and $\theta^*_{V\text{-max}}$ Were Obtained by Using $\underline{\theta} = -4.00$ and $\bar{\theta} = 4.00$.

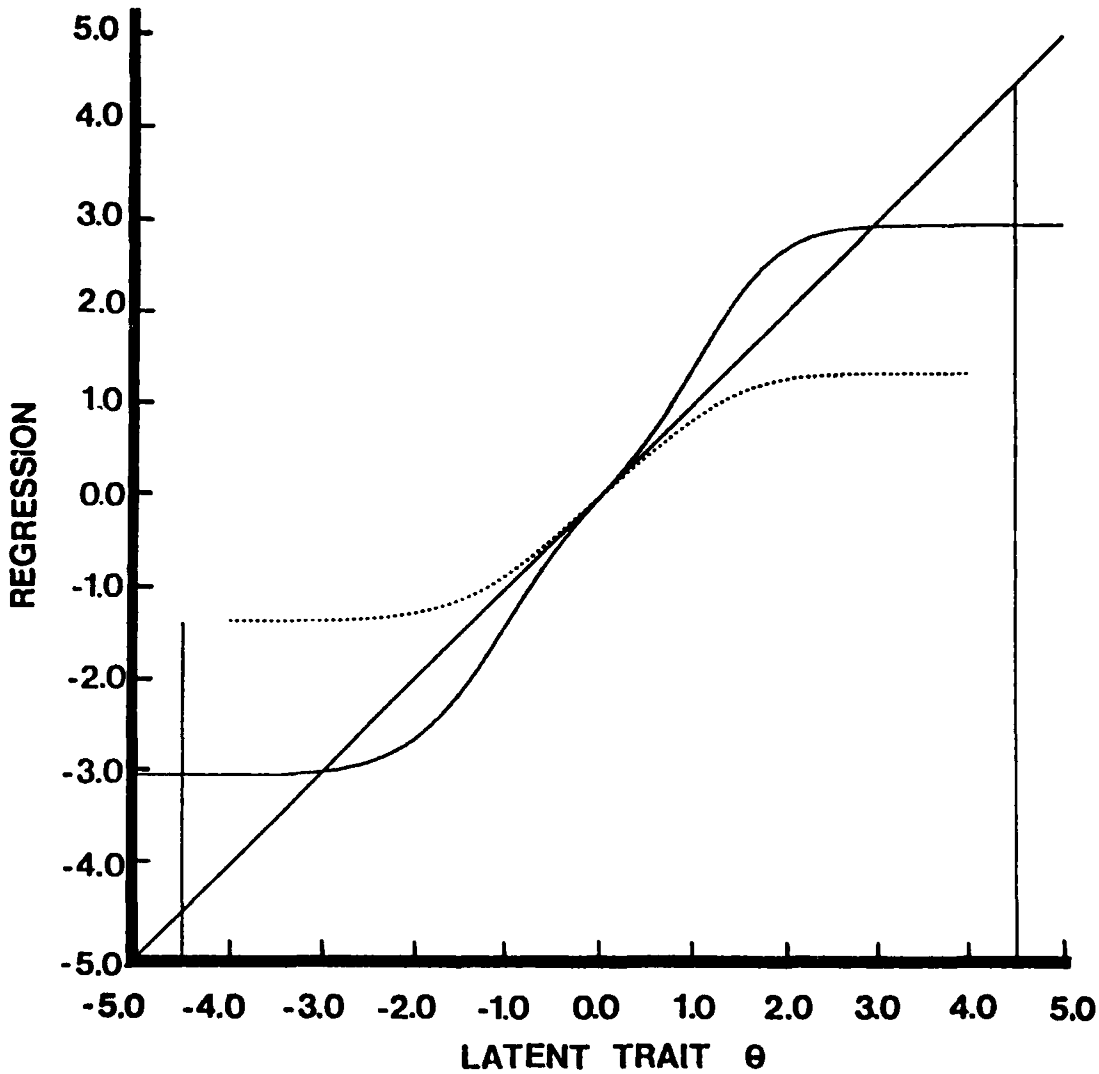

FIGURE A-1 (Continued)

$\theta^*_{V\text{-min}}$ and $\theta^*_{V\text{-max}}$ Were Obtained by Using $\underline{\theta} = -4.50$ and $\bar{\theta} = 4.50$.

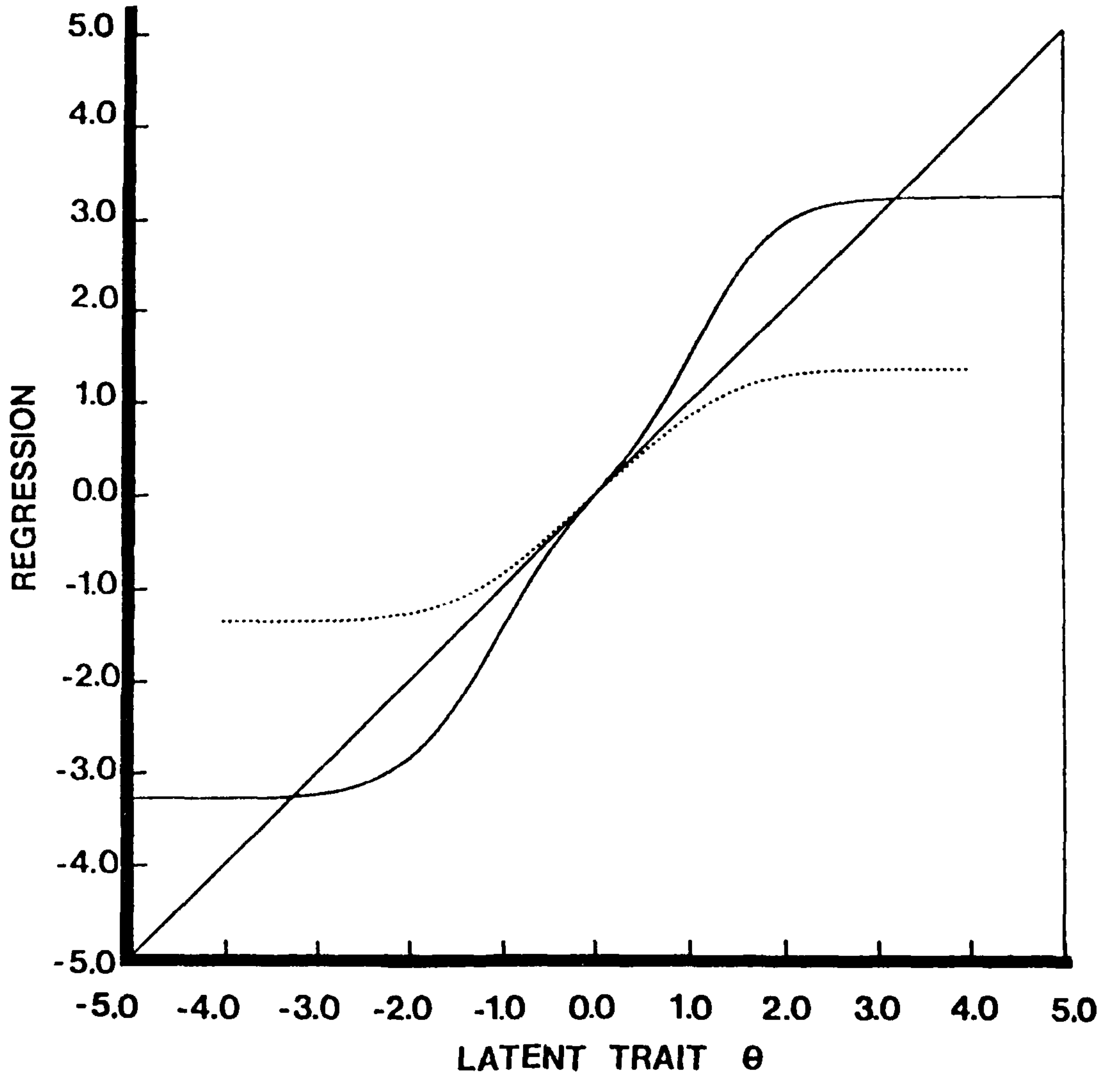

FIGURE A-1 (Continued)

$\theta^*_{V\text{-min}}$ and $\theta^*_{V\text{-max}}$ Were Obtained by Using $\underline{\theta} = -5.00$ and $\bar{\theta} = 5.00$.

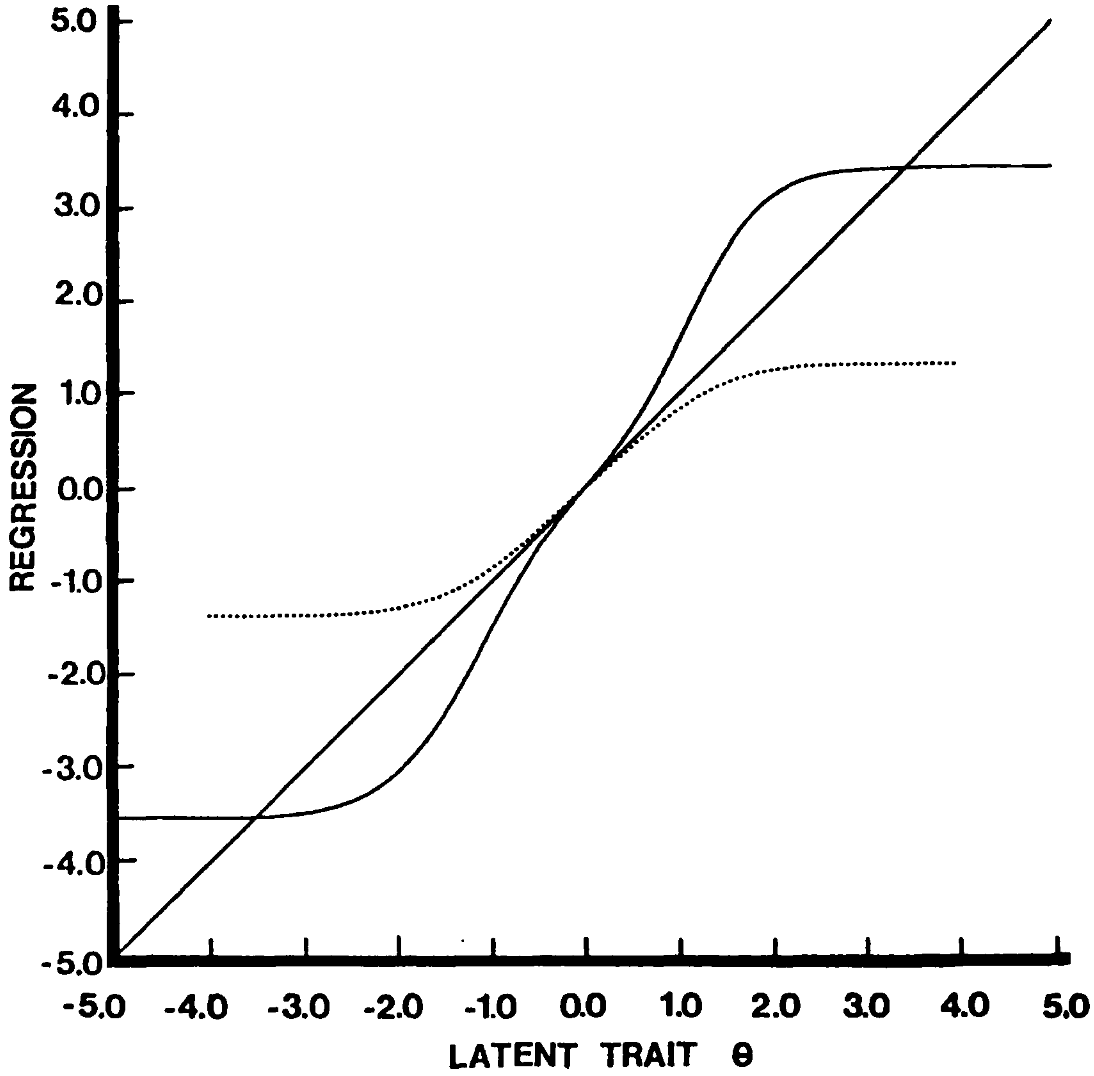

FIGURE A-1 (Continued)

θ^*_{V-min} and θ^*_{V-max} Were Obtained by Using $\underline{\theta} = -5.50$ and $\bar{\theta} = 5.50$.

TABLE A-5

Maximum Likelihood Estimates for the 126 Possible Response Patterns on LIS-U, Excluding the Two Extreme Response Patterns, in Which All the Answers Are Incorrect, and All of Them Are Correct, Respectively.

	Response Pattern	MLE	Number of Iterations
2	2111111	-1.2260	6
3	1211111	-0.8567	5
4	2211111	-0.4968	4
5	1121111	-1.1661	6
6	2121111	-0.7267	5
7	1221111	-0.5030	4
8	2221111	-0.1602	3
9	1112111	-1.3167	6
10	2112111	-0.8129	6
11	1212111	-0.5726	5
12	2212111	-0.2307	4
13	1122111	-0.7983	5
14	2122111	-0.4491	5
15	1222111	-0.2520	4
16	2222111	0.0997	4
17	1111211	-1.0350	6
18	2111211	-0.6555	5
19	1211211	-0.4522	4
20	2211211	-0.1251	4
21	1121211	-0.6523	5
22	2121211	-0.3400	5
23	1221211	-0.1515	4
24	2221211	0.2005	4
25	1112211	-0.7231	5
26	2112211	-0.4021	5
27	1212211	-0.2151	4
28	2212211	0.1241	4
29	1122211	-0.4125	5
30	2122211	-0.1188	4
31	1222211	0.0832	4
32	2222211	0.4735	5
33	1111121	-0.7831	5
34	2111121	-0.4382	5
35	1211121	-0.3170	4
36	2211121	-0.0147	3
37	1121121	-0.4937	5
38	2121121	-0.2220	4
39	1221121	-0.0448	4
40	2221121	0.2950	5
41	1112121	-0.5503	5

Note: In this table, 1 is used instead of 0 for the incorrect answers, and 2 is used instead of 1 for the correct answers.

TABLE A-5 (Continued)

	Response Pattern	MLE	Number of Iterations
42	2112121	-0.2767	4
43	1212121	-0.1045	4
44	2212121	0.2192	4
45	1122121	-0.2915	4
46	2122121	-0.0240	4
47	1222121	0.1752	4
48	2222121	0.5636	5
49	1111221	-0.4525	5
50	2111221	-0.1923	4
51	1211221	-0.0196	4
52	2211221	0.3116	5
53	1121221	-0.2102	4
54	2121221	0.0524	4
55	1221221	0.2618	4
56	2221221	0.6273	5
57	1112221	-0.2605	4
58	2112221	-0.0018	3
59	1212221	0.1938	4
60	2212221	0.5732	5
61	1122221	-0.0275	4
62	2122221	0.2443	4
63	1222221	0.5030	5
64	2222221	1.0301	6
65	1111112	-0.7214	5
66	2111112	-0.4411	5
67	1211112	-0.2772	4
68	2211112	0.0138	3
69	1121112	-0.4485	5
70	2121112	-0.1881	4
71	1221112	-0.0164	3
72	2221112	0.3085	5
73	1112112	-0.5022	5
74	2112112	-0.2404	4
75	1212112	-0.0739	4
76	2212112	0.2364	4
77	1122112	-0.2562	4
78	2122112	0.0008	3
79	1222112	0.1936	4
80	2222112	0.5608	5
81	1111212	-0.4110	5
82	2111212	-0.1609	4
83	1211212	0.0068	3
84	2211212	0.3238	5

Note: In this table, 1 is used instead of 0 for the incorrect answers, and 2 is used instead of 1 for the correct answers.

TABLE A-5 (Continued)

	Response Pattern	MLE	Number of Iterations
85	1121212	-0.1794	4
86	2121212	0.0734	4
87	1221212	0.2759	4
88	2221212	0.6661	5
89	1112212	-0.2277	4
90	2112212	0.0713	4
91	1212212	0.2108	4
92	2212212	0.5698	5
93	1122212	-0.0043	3
94	2122212	0.2571	4
95	1222212	0.5042	5
96	2222212	0.9851	5
97	1111122	-0.3069	4
98	2111122	-0.0753	4
99	1211122	0.0896	4
100	2211122	0.4029	5
101	1121122	-0.0958	4
102	2121122	0.1455	4
103	1221122	0.3518	5
104	2221122	0.7529	5
105	1112122	-0.1410	4
106	2112122	0.0950	3
107	1212122	0.2863	5
108	2212122	0.6501	5
109	1122122	0.0679	3
110	2122122	0.3222	4
111	1222122	0.5795	5
112	2222122	1.0985	6
113	1111222	-0.0751	4
114	2111222	0.1608	4
115	1211222	0.3644	5
116	2211222	0.7577	5
117	1121222	0.1313	4
118	2121222	0.3906	4
119	1221222	0.6758	5
120	2221222	1.3028	6
121	1112222	0.0842	4
122	2112222	0.3334	4
123	1212222	0.5873	5
124	2212222	1.0957	6
125	1122222	0.2963	4
126	2122222	0.5807	5
127	1222222	0.9714	6

Note; In this table, 1 is used instead of 0 for the incorrect answers, and 2 is used instead of 1 for the correct answers.

DISTRIBUTION LIST

Navy

1 Dr. Jack R. Borsting
Provost & Academic Dean
U.S. Naval Postgraduate School
Monterey, CA 93940

1 Dr. Robert Breaux
Code N-711
NAVTRAEQUIPCEN
Orlando, FL 32813

1 COMNAVMILPERSCOM (N-6C)
Dept. of Navy
Washington, DC 20370

1 Dr. Larry Dean, LT, MSC, USN
Psychology Department
Naval Submarine Medical Research Lab
Naval Submarine Base
Groton, CT 06340

1 Dr. Richard Elster
Department of Administrative Sciences
Naval Postgraduate School
Monterey, CA 93940

1 DR. PAT FEDERICO
NAVY PERSONNEL R&D CENTER
SAN DIEGO, CA 92152

1 Mr. Paul Foley
Navy Personnel R&D Center
San Diego, CA 92152

1 Dr. John Ford
Navy Personnel R&D Center
San Diego, CA 92152

1 Dr. Patrick R. Harrison
Psychology Course Director
LEADERSHIP & LAW DEPT. (7b)
DIV. OF PROFESSIONAL DEVELOPMMENT
U.S. NAVAL ACADEMY
ANNAPOLIS, MD 21402

1 Dr. Norman J. Kerr
Chief of Naval Technical Training
Naval Air Station Memphis (75)
Millington, TN 38054

Navy

1 Dr. William L. Maloy
Principal Civilian Advisor for
Education and Training
Naval Training Command, Code 00A
Pensacola, FL 32508

1 Dr. Kneale Marshall
Scientific Advisor to DCNO(MPT)
OP01T
Washington DC 20370

1 CAPT Richard L. Martin, USN
Prospective Commanding Officer
USS Carl Vinson (CVN-70)
Newport News Shipbuilding and Drydock Co
Newport News, VA 23607

1 Dr. James McBride
Navy Personnel R&D Center
San Diego, CA 92152

1 Dr. George Moeller
Head, Human Factors Dept.
Naval Submarine Medical Research Lab
Groton, CN 06340

1 Library
Naval Health Research Center
P. O. Box 85122
San Diego, CA 92138

1 Naval Medical R&D Command
Code 44
National Naval Medical Center
Betheada, MD 20014

1 Dr. Ted M. I. Yellen
Technical Information Office, Code 201
NAVY PERSONNEL R&D CENTER
SAN DIEGO, CA 92152

1 Library, Code P201L
Navy Personnel R&D Center
San Diego, CA 92152

1 Technical Director
Navy Personnel R&D Center
San Diego, CA 92152

Navy

6 Commanding Officer
Naval Research Laboratory
Code 2627
Washington, DC 20390

1 Psychologist
ONR Branch Office
Bldg 114, Section D
666 Summer Street
Boston, MA 02210

1 Psychologist
ONR Branch Office
536 S. Clark Street
Chicago, IL 60605

1 Office of Naval Research
Code 437
800 N. Quincy SStreet
Arlington, VA 22217

5 Personnel & Training Research Programs
(Code 458)
Office of Naval Research
Arlington, VA 22217

1 Psychologist
ONR Branch Office
1030 East Green Street
Pasadena, CA 91101

1 Office of the Chief of Naval Operations
Research, Development, and Studies Branc
(OP-102)
Washington, DC 20350

1 Captain Donald F. Parker, USN
Commanding Officer
Navy Personnel R&D Center
San Diego, CA 92152

1 LT Frank C. Petho, MSC, USN (Ph.D)
Code L51
Naval Aerospace Medical Research Laborat
Pensacola, FL 32508

Navy

1 Director, Research & Analysis Divisic
Plans and Policy Department
Navy Recruiting Command
4015 Wilson Boulevard
Arlington, VA 22203

1 Dr. Bernard Rimland (03B)
Navy Personnel R&D Center
San Diego, CA 92152

1 Mr. Arnold Rubenstein
Naval Personnel Support Technology
Naval Material Command (08T244)
Room 1044, Crystal Plaza #5
2221 Jefferson Davis Highway
Arlington, VA 20360

1 Dr. Worth Scanland
Chief of Naval Education and Training
Code N-5
NAS, Pensacola, FL 32508

1 Dr. Robert G. Smith
Office of Chief of Naval Operations
OP-987H
Washington, DC 20350

1 Dr. Alfred F. Smode
Training Analysis & Evaluation Group
(TAEG)
Dept. of the Navy
Orlando, FL 32813

1 Dr. Richard Sorensen
Navy Personnel R&D Center
San Diego, CA 92152

1 Dr. Ronald Weitzman
Code 54 WZ
Department of Administrative Sciences
U. S. Naval Postgraduate School
Monterey, CA 93940

1 Dr. Robert Wisher
Code 309
Navy Personnel R&D Center
San Diego, CA 92152

Navy

1 DR. MARTIN F. WISKOFF
NAVY PERSONNEL R& D CENTER
SAN DIEGO, CA 92152

1 Mr John H. Wolfe
Code P310
U. S. Navy Personnel Research and
Development Center
San Diego, CA 92152

Army

1 Technical Director
U. S. Army Research Institute for the
Behavioral and Social Sciences
5001 Eisenhower Avenue
Alexandria, VA 22333

1 HQ USAREUE & 7th Army
ODCSOPS
USAAREUE Director of GED
APO New York 09403

1 Col Gary W. Bloedorn
US Army TRADOC Systems Analysis Activity
Attn: ATAA-TH
WSMR, NM 88002

1 DR. RALPH DUSEK
U.S. ARMY RESEARCH INSTITUTE
5001 EISENHOWER AVENUE
ALEXANDRIA, VA 22333

1 Dr. Myron Fischl
U.S. Army Research Institute for the
Social and Behavioral Sciences
5001 Eisenhower Avenue
Alexandria, VA 22333

1 Dr. Milton S. Katz
Training Technical Area
U.S. Army Research Institute
5001 Eisenhower Avenue
Alexandria, VA 22333

1 Dr. Harold F. O'Neil, Jr.
Attn: PERI-OK
Army Research Institute
5001 Eisenhower Avenue
Alexandria, VA 22333

1 LTC Michael Plummer
Chief, Leadership & Organizational
Effectiveness Division
Office of the Deputy Chief of Staff
for Personnel
Dept. of the Army
Pentagon, Washington DC 20301

Army

1 DR. JAMES L. RANEY
U.S. ARMY RESEARCH INSTITUTE
5001 EISENHOWER AVENUE
ALEXANDRIA, VA 22333

1 Mr. Robert Ross
U.S. Army Research Institute for the Social and Behavioral Sciences
5001 Eisenhower Avenue
Alexandria, VA 22333

1 Dr. Robert Sasmor
U. S. Army Research Institute for the Behavioral and Social Sciences
5001 Eisenhower Avenue
Alexandria, VA 22333

1 Commandant
US Army Institute of Administration
Attn: Dr. Sherrill
FT Benjamin Harrison, IN 46256

1 Dr. Frederick Steinheiser
U. S. Army Reserch Institute
5001 Eisenhower Avenue
Alexandria, VA 22333

1 Dr. Joseph Ward
U.S. Army Research Institute
5001 Eisenhower Avenue
Alexandria, VA 22333

Air Force

1 Air Force Human Resources Lab
AFHRL/MPD
Brooks AFB, TX 78235

1 U.S. Air Force Office of Scientific Research
Life Sciences Directorate, NL
Bolling Air Force Base
Washington, DC 20332

1 Air University Library
AUL/LSE 76/443
Maxwell AFB, AL 36112

1 Dr. Earl A. Alluisi
HQ, AFHRL (AFSC)
Brooks AFB, TX 78235

1 Dr. Genevieve Haddad
Program Manager
Life Sciences Directorate
AFOSR
Bolling AFB, DC 20332

1 Research and Measurment Division
Research Branch, AFMPC/MPCYPR
Randolph AFB, TX 78148

1 Dr. Malcolm Ree
AFHRL/MP
Brooks AFB, TX 78235

1 Dr. Marty Rockway (AFHRL/TT)
Lowry AFB
Colorado 80230

1 Dr. Frank Schufletowski
U.S. Air Force
ATC/XPTD
Randolph AFB, TX 78148

1 Jack A. Thorpe, Maj., USAF
Naval War College
Providence, RI 02846

1 Dr. Joe Ward, Jr.
AFHRL/MPMD
Brooks AFB, TX 78235

Marines

1 Dr. H. William Greenup
Education Advisor (E031)
Education Center, MCDEC
Quantico, VA 22134

1 Major Howard Langdon
Headquarters, Marine Corps
OTTI 31
Arlington Annex
Columbia Pike at Arlington Ridge Rd.
Arlington, VA 20380

1 Director, Office of Manpower Utilization
HQ, Marine Corps (MPU)
BCB, Bldg. 2009
Quantico, VA 22134

1 Headquarters, U. S. Marine Corps
Code MPI-20
Washington, DC 20380

1 Special Assistant for Marine
Corps Matters
Code 100M
Office of Naval Research
800 N. Quincy St.
Arlington, VA 22217

1 Major Michael L. Patrow, USMC
Headquarters, Marine Corps
(Code MPI-20)
Washington, DC 20380

1 DR. A.L. SLAFKOSKY
SCIENTIFIC ADVISOR (CODE RD-1)
HQ, U.S. MARINE CORPS
WASHINGTON, DC 20380

CoastGuard

1 Chief, Psychological Reserch Branch
U. S. Coast Guard (G-P-1/2/TP42)
Washington, DC 20593

1 Mr. Thomas A. Warm
U. S. Coast Guard Institute
P. O. Substation 18
Oklahoma City, OK 73169

Other DoD

12 Defense Documentation Center
Cameron Station, Bldg. 5
Alexandria, VA 22314
Attn: TC

1 Dr. Dexter Fletcher
ADVANCED RESEARCH PROJECTS AGENCY
1400 WILSON BLVD.
ARLINGTON, VA 22209

1 Dr. William Graham
Testing Directorate
MEPCOM/MEPCT-P
Ft. Sheridan, IL 60037

1 Director, Research and Data
OASD(MRA&L)
3B919, The Pentagon
Washington, DC 20301

1 Military Assistant for Training and
Personnel Technology
Office of the Under Secretary of Defense
for Research & Engineering
Room 3D129, The Pentagon
Washington, DC 20301

1 MAJOR Wayne Sellman, USAF
Office of the Assistant Secretary
of Defense (MRA&L)
3B930 The Pentagon
Washington, DC 20301

Civil Govt

1 Dr. Susan Chipman
Learning and Development
National Institute of Education
1200 19th Street NW
Washington, DC 20208

1 Dr. Lorraine D. Eyde
Personnel R&D Center
Office of Personnel Management of USA
1900 EStreet NW
Washington, D.C. 20415

1 Dr. Jerry Lehnus
REGIONAL PSYCHOLOGIST
U.S. Office of Personnel Management
230 S. DEARBORN STREET
CHICAGO, IL 60604

1 Dr. Andrew R. Molnar
Science Education Dev.
and Research
National Science Foundation
Washington, DC 20550

1 Personnel R&D Center
Office of Personnel Managment
1900 E Street NW
Washington, DC 20415

1 Dr. H. Wallace Sinaiko
Program Director
Manpower Research and Advisory Services
Smithsonian Institution
801 North Pitt Street
Alexandria, VA 22314

1 Dr. Vern W. Urry
Personnel R&D Center
Office of Personnel Management
1900 E Street NW
Washington, DC 20415

1 Dr. Joseph L. Young, Director
Memory & Cognitive Processes
National Science Foundation
Washington, DC 20550

Non Govt

1 Dr. Erling B. Andersen
Department of Statistics
Studiestraede 6
1455 Copenhagen
DENMARK

1 1 psychological research unit
Dept. of Defense (Army Office)
Campbell Park Offices
Canberra ACT 2600, Australia

1 Dr. Jackson Beatty
Department of Psychology
University of California
Los Angeles, CA 90024

1 Dr. Isaac Bejar
Educational Testing Service
Princeton, NJ 08450

1 Dr. John Bergan
School of Education
University of Arizona
Tuscon AZ 85721

1 Dr. Werner Birke
DezWPs im Streitkraefteamt
Postfach 20 50 03
D-5300 Bonn 2
WEST GERMANY

1 Dr. R. Darrel Bock
Department of Education
University of Chicago
Chicago, IL 60637

1 Dr. Nicholas A. Bond
Dept. of Psychology
Sacramento State College
600 Jay Street
Sacramento, CA 95819

1 Dr. Robert Brennan
American College Testing Programs
P. O. Box 168
Iowa City, IA 52240

Non Govt

1 DR. C. VICTOR BUNDERSON
WICAT INC.
UNIVERSITY PLAZA, SUITE 10
1160 SO. STATE ST.
OREM, UT 84057

1 Dr. Anthony Cancelli
School of Education
University of Arizona
Tuscon, AZ 85721

1 Dr. John B. Carroll
Psychometric Lab
Univ. of No. Carolina
Davie Hall 013A
Chapel Hill, NC 27514

1 Charles Myers Library
Livingstone House
Livingstone Road
Stratford
London E15 2LJ
ENGLAND

1 Dr. Kenneth E. Clark
College of Arts & Sciences
University of Rochester
River Campus Station
Rochester, NY 14627

1 Dr. Norman Cliff
Dept. of Psychology
Univ. of So. California
University Park
Los Angeles, CA 90007

1 Dr. William E. Coffman
Director, Iowa Testing Programs
334 Lindquist Center
University of Iowa
Iowa City, IA 52242

1 Dr. Meredith P. Crawford
American Psychological Association
1200 17th Street, N.W.
Washington, DC 20036

Non Govt

1 Dr. Hans Crombag
Education Research Center
University of Leyden
Boerhaavelaan 2
2334 EN Leyden
The NETHERLANDS

1 Director
Behavioural Sciences Division
Defence & Civil Institute of
Environmental Medicine
Post Office Box 2000
Downsview, Ontario M3M 3B9
CANADA

1 LCOL J. C. Eggenberger
DIRECTORATE OF PERSONNEL APPLIED RESEARC
NATIONAL DEFENCE HQ
101 COLONEL BY DRIVE
OTTAWA, CANADA K1A 0K2

1 ERIC Facility-Acquisitions
4833 Rugby Avenue
Bethesda, MD 20014

1 Dr. Leonard Feldt
Lindquist Center for Measurment
University of Iowa
Iowa City, IA 52242

1 Dr. Richard L. Ferguson
The American College Testing Program
P.O. Box 168
Iowa City, IA 52240

1 Dr. Victor Fields
Dept. of Psychology
Montgomery College
Rockville, MD 20850

1 Univ. Prof. Dr. Gerhard Fischer
Liebiggasse 5/3
A 1010 Vienna
AUSTRIA

1 Professor Donald Fitzgerald
University of New England
Armidale, New South Wales 2351
AUSTRALIA

Non Govt

1 Dr. John R. Frederiksen
Bolt Beranek & Newman
50 Moulton Street
Cambridge, MA 02138

1 DR. ROBERT GLASER
LRDC
UNIVERSITY OF PITTSBURGH
3939 O'HARA STREET
PITTSBURGH, PA 15213

1 DR. JAMES G. GREENO
LRDC
UNIVERSITY OF PITTSBURGH
3939 O'HARA STREET
PITTSBURGH, PA 15213

1 Dr. Ron Hambleton
School of Education
University of Massechusetts
Amherst, MA 01002

1 Dr. Chester Harris
School of Education
University of California
Santa Barbara, CA 93106

1 Dr. Frederick Hayes-Roth
The Rand Corporation
1700 Main Street
Santa Monica, CA 90406

1 Dr. Lloyd Humphreys
Department of Psychology
University of Illinois
Champaign, IL 61820

1 Library
HumRRO/Western Division
27857 Berwick Drive
Carmel, CA 93921

1 Dr. Steven Hunka
Department of Education
University of Alberta
Edmonton, Alberta
CANADA

Non Govt

1 Dr. Earl Hunt
Dept. of Psychology
University of Washington
Seattle, WA 98105

1 Dr. Huynh Huynh
College of Education
University of South Carolina
Columbia, SC 29208

1 Dr. Douglas H. Jones
Rm T-255
Educational Testing Service
Princeton, NJ 08450

3 Journal Supplement Abstract Service
American Psychological Association
1200 17th Street N.W.
Washington, DC 20036

1 Professor John A. Keats
University of Newcastle
AUSTRALIA 2308

1 Dr. Stephen Kosslyn
Harvard University
Department of Psychology
33 Kirkland Street
Cambridge, MA 02138

1 Mr. Marlin Kroger
1117 Via Goleta
Palos Verdes Estates, CA 90274

1 Dr. Alan Lesgold
Learning R&D Center
University of Pittsburgh
Pittsburgh, PA 15260

1 Dr. Michael Levine
210 Education Building
University of Illinois
Champaign, IL 61820

Non Govt

1 Dr. Charles Lewis
Faculteit Sociale Wetenschappen
Rijksuniversiteit Groningen
Oude Boteringestraat
Groningen
NETHERLANDS

1 Dr. Robert Linn
College of Education
University of Illinois
Urbana, IL 61801

1 Dr. Frederick M. Lord
Educational Testing Service
Princeton, NJ 08540

1 Dr. James Lumsden
Department of Psychology
University of Western Australia
Nedlands W.A. 6009
AUSTRALIA

1 Dr. Gary Marco
Educational Testing Service
Princeton, NJ 08450

1 Dr. Scott Maxwell
Department of Psychology
University of Houston
Houston, TX 77004

1 Dr. Samuel T. Mayo
Loyola University of Chicago
820 North Michigan Avenue
Chicago, IL 60611

1 Professor Jason Millman
Department of Education
Stone Hall
Cornell University
Ithaca, NY 14853

1 Dr. Melvin R. Novick
356 Lindquist Center for Measurment
University of Iowa
Iowa City, IA 52242

Non Govt

1 Dr. Jesse Orlansky
Institute for Defense Analyses
400 Army Navy Drive
Arlington, VA 22202

1 Dr. James A. Paulson
Portland State University
P.O. Box 751
Portland, OR 97207

1 MR. LUIGI PETRULLO
2431 N. EDGEWOOD STREET
ARLINGTON, VA 22207

1 DR. DIANE M. RAMSEY-KLEE
R-K RESEARCH & SYSTEM DESIGN
3947 RIDGEMONT DRIVE
MALIBU, CA 90265

1 MINRAT M. L. RAUCH
P II 4
BUNDESMINISTERIUM DER VERTEIDIGUNG
POSTFACH 1328
D-53 BONN 1, GERMANY

1 Dr. Mark D. Reckase
Educational Psychology Dept.
University of Missouri-Columbia
4 Hill Hall
Columbia, MO 65211

1 Dr. Andrew M. Rose
American Institutes for Research
1055 Thomas Jefferson St. NW
Washington, DC 20007

1 Dr. Leonard L. Rosenbaum, Chairman
Department of Psychology
Montgomery College
Rockville, MD 20850

1 Dr. Lawrence Rudner
403 Elm Avenue
Takoma Park, MD 20012

1 Dr. J. Ryan
Department of Education
University of South Carolina
Columbia, SC 29208

Non Govt

1 DR. WALTER SCHNEIDER
DEPT. OF PSYCHOLOGY
UNIVERSITY OF ILLINOIS
CHAMPAIGN, IL 61820

1 DR. ROBERT J. SEIDEL
INSTRUCTIONAL TECHNOLOGY GROUP
HUMRRO
300 N. WASHINGTON ST.
ALEXANDRIA, VA 22314

1 Dr. Kazuo Shigemasu
University of Tohoku
Department of Educational Psychology
Kawauchi, Sendai 980
JAPAN

1 Dr. Edwin Shirkey
Department of Psychology
University of Central Florida
Orlando, FL 32816

1 Dr. Robert Smith
Department of Computer Science
Rutgers University
New Brunswick, NJ 08903

1 Dr. Richard Snow
School of Education
Stanford University
Stanford, CA 94305

1 Dr. Kathryn T. Spoehr
Department of Psychology
Brown University
Providence, RI 02912

1 Dr. Robert Sternberg
Dept. of Psychology
Yale University
Box 11A, Yale Station
New Haven, CT 06520

1 Dr. David Stone
ED 236
SUNY, Albany
Albany, NY 12222

Non Govt

1 DR. PATRICK SUPPES
INSTITUTE FOR MATHEMATICAL STUDIES IN THE SOCIAL SCIENCES
STANFORD UNIVERSITY
STANFORD, CA 94305

1 Dr. Hariharan Swaminathan
Laboratory of Psychometric and Evaluation Research
School of Education
University of Massachusetts
Amherst, MA 01003

1 Dr. Brad Sympson
Psychometric Research Group
Educational Testing Service
Princeton, NJ 08541

1 Dr. Kikumi Tatsuoka
Computer Based Education Research Laboratory
252 Engineering Research Laboratory
University of Illinois
Urbana, IL 61801

1 Dr. David Thissen
Department of Psychology
University of Kansas
Lawrence, KS 66044

1 Dr. Douglas Towne
Univ. of So. California
Behavioral Technology Labs
1845 S. Elena Ave.
Redondo Beach, CA 90277

1 Dr. J. Uhlaner
Perceptronics, Inc.
6271 Variel Avenue
Woodland Hills, CA 91364

1 Dr. Howard Wainer
Bureau of Social SCience Research
1990 M Street, N. W.
Washington, DC 20036

Non Govt

1 Dr. Phyllis Weaver
Graduate School of Education
Harvard University
200 Larsen Hall, Appian Way
Cambridge, MA 02138

1 Dr. David J. Weiss
N660 Elliott Hall
University of Minnesota
75 E. River Road
Minneapolis, MN 55455

1 DR. SUSAN E. WHITELY
PSYCHOLOGY DEPARTMENT
UNIVERSITY OF KANSAS
LAWRENCE, KANSAS 66044

1 Dr. Wolfgang Wildgrube
Streitkraefteamt
Box 20 50 03
D-5300 Bonn 2
WEST GERMANY

1 Dr. J. Arthur Woodward
Department of Psychology
University of California
Los Angeles, CA 90024

Navy

1 Mr. Donald Calder
Office of Naval Research
325 Hinman Research Building
Atlanta, GA 30332

Army

1 Dr. Randall M. Chambers, Ph.D.
U.S. Army Research Institute
for the Behavioral and Social
Sciences
Fort Sill Field Unit
P. O. Box 3066
Fort Sill, OK 73503

Non Govt

1 Dr. Bert F. Green
Department of Psychology
The John's Hopkins University
Charles at 34th Street
Baltimore, MD 21218

1 Dr. Ron Hambleton
School of Education
University of Massachusetts
Amherst, Mass. 01002

1 Dr. William W. Turnbull
Educational Testing Service
Princeton, NJ 08540

1 Mr. Philip S. Livingston
1500 Sparkman #32CC
Huntsville, AL 35805

Non Govt

1 Dr. Isaac I. Bejar
Department of Psychology
Elliott Hall
75 East River Road
Minneapolis, Minnesota 55455

1 Dr. George Woods
1106 Newport Ave.
Victoria, B. C.
V8S 5E4 Canada

1 Dr. P. Mengal
Faculte' de Psychologie
et des Sciences de l'Education
Universite' de Geneve
3 fl. de l'Universite
1201 Geneva SWITZERLAND

1 Dr. Wim J. van der Linden
Vakgroep Onderwisjskunde
Postbus 217
7500 EA Enschede
The Netherlands

1 Dr. Lowell Schipper
Department of Psychology
Bowling Green State University
Bowling Green, Ohio 43403

1 Dr. Lutz Hornke
University Duesseldorf
Erz. Wiss.
D-4000 Duesseldorf
WEST GERMANY

1 Dr. Wolfgang Buchtala
8346 Simbach Inn
Postfach 1306
Industriestrasse 1
WEST GERMANY